tanulmányok

80/1978

MTA Számitástechnikai és Automatizálási Kutató Intézet Budapest

MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMITÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE

STUDIES IN APPLIED STOCHASTIC PROGRAMMING

I.

(Cikkgyüjtemény)

Edited
by

ANDRÁS PRÉKOPA

Tanulmányok 80/1978.

A kiadásért felelős:

DR VÁMOS TIBOR

ISBN 963 311 065 3

ISSN 0324-2951

Készült a
KSH Nemzetközi Számitástechnikai Oktató és Tájékoztató
Központ Reprográfiai Üzemében

CONTENTS

THE STABIL STOCHASTIC PROGRAMMING MODEL AND ITS EXPERIMENTAL APPLICATION TO THE ELECTRICAL ENERGY SECTOR OF THE HUNGARIAN ECONOMY

András Prékopa
Budapest

Sándor Ganczer
Budapest

István Deák
Budapest

Károly Patyi
Budapest

ABSTRACT

The stochastic programming model named STABIL is the model (1.1), where we minimize a linear or nonlinear objective function under a probabilistic and some further constrains. In this paper we are concerned with a special case of this model-type, where the functions $g_1, \ldots, g_{m+M}, f$ are linear. The model is applied to the 4th Five-Year Plan of the electrical energy sector of the Hungarian economy where the underlying deterministic model is the corresponding deterministic sector model of the 4th Five-Year Plan. This application is of experimental type, as the construction of the model and the computations were performed in the time, in which the 4th Five-Year Plan was already running. In this paper we describe the model, the solving algorithm, the computer program, the parameters of the model and the computational results.

Very interesting phaenomenon is that that there is no significant difference between the optimum value of the deterministic underlying problem and that of the related STABIL model, but different optimal solutions turned out. As regards the reliability level of the optimal solution of the deterministic underlying problem it turned out to be very low while that of the optimal solution of the STABIL model is high. The appearance of this phaenomenon makes interesting the numerical example also in itself independently from

the special problem in connection of which it occured.

1. Introduction.

In this paper we describe in detail the probabilistic constrained stochastic programming model named STABIL. The model, its theory and the solving algorithm are described in a shorther form in the papers [4],[5] and [8] . Further we apply the model to the electrical energy sector of the Hungarian economy in the frame of the 4th Five-Year Plan. The description of the linear programming model of the 4th Five-Year Plan can be found in [1] and [3].

The STABIL name of the model is introduced in this paper. This name is not an abbreviation. The reason why we have chosen it is that the model contains a probabilistic constraint, where we prescribe a high probability level with which the system has to operate.

In the paper the computer program system of the solving algorithm of the model (1.1) is also described in a short form. For further details the reader is referred to the paper [2].

The applications of our model is experimental for two reasons. Firstly because the underlying problem of the stochastic programming model was the deterministic electrical energy sector model of the 4th Five-Year Plan already going on in the time in which this work was done thus we could not think of the practical application of the numerical results. Secondly, there were lacks of information connected with the relevant probability distribution of the occuring random variables and some subjective considerations were applied in order to fix this distribution.

The STABIL stochastic progarmming model is the following

$$
\begin{aligned}
& G(\underline{x}) = P(g_i(\underline{x}) \geqq \xi_i, && i = 1, \ldots, m) \geqq p, \\
& && g_i(\underline{x}) \geq b_i, && i = m+1, \ldots, m+M, \\
& && \min f(\underline{x}) \ . &&
\end{aligned}
\tag{1.1}
$$

We are interested in that special case of the model where the functions $g_1, \ldots, g_{m+M}, f$ are linear. The joint distribution of the random variables $\xi_1, \ldots, \xi_m$ is supposed to be continuous with logarithmic concave joint density. The case we are dealing with here is the non-degenerated multivariate normal distribution which has the mentioned property. In the model (1.1) p is a prescribed probability, $0 < p < 1$, the choice of which depends on us. p is of course near unity in practice. In the electrical energy problem the value of p was 0.9 respectively 0.95. As regards the further characteristics of the problem, $m=4$, $M=106$, where the individual upper and lower bounds and the nonegativity constraints are incorporated and the vector $\underline{x}$ has 46 components.

The deterministic model, which is the starting point of the stochastic programming model construction i.e. the underlying deterministic model of the model (1.1) is the following:

$$
\begin{aligned}
& g_i(\underline{x}) \geq b_i, \quad i=1,\ldots,m \\
& g_i(\underline{x}) \geq b_i, \quad i=m+1,\ldots,m+M, \\
& \min f(\underline{x})
\end{aligned}
\tag{1.2}
$$

The values on the right hand side of the first row in (1.2) were supposed to be random and thus we were lead to model (1.1). In the practical model constructions the random variables are frequently replaced by their expectations. In such cases if we formulate stochastic programming decision model and look for the probability distribution of the random variables appearing in the model, then the expectations can be taken out from the underlying deterministic model. Following this methodology, the random variables of the model (1.1) will be written in the form

$$b_i + \sigma_i \beta_i, \quad i=1,\ldots,m,$$

where

$$E(\beta_i) = 0, \quad i=1,\ldots,m. \tag{1.3}$$

We may also suppose that

$$D(\beta_i) = 1, \qquad i=1,\ldots,m. \tag{1.4}$$

The symbols E,D are used to denote the expectation and the dispersion, respectively. As we already mentioned in the special case of (1.1) we are interested in, the functions $g_1,\ldots,g_{m+m},f$ are linear. Let us introduce the notations

$$\begin{aligned} g_i(\underline{x}) &= \underline{a}_i'\underline{x}, \qquad i=1,\ldots,m+M, \\ f(\underline{x}) &= \underline{c}'\underline{x}. \end{aligned} \tag{1.5}$$

The mathematical model and the solving algorithms presented in this paper were constructed by A. Prékopa. The evaluation procedure of the distribution function of the multivariate normal distribution, the program system for the solution of the problem (except for the linear programming package which was worked out by G. Kéri), and the numerical computations were done by I. Deák. The formulation of the deterministic underlying problem and all data collections were done by S. Ganczer and K. Patyi. The application of the stochastic programming model to the electrical energy sector i.e. the model adaptation of the model was done jointly by the authors of the present paper.

2. Detailed description of the stochastic programming model

We are concerned with the numerical solution of the following special case of the STABIL model:

$$\begin{aligned} G(\underline{x}) = P\left(\frac{1}{\sigma_i}(\underline{a}_i'\underline{x}-b_i) \geq \beta_i, \quad i=1,\ldots,m\right) \geq p, \\ \underline{a}_i'\underline{x} \geq b_i \qquad\qquad i=m+1,\ldots,m+M, \\ \min \underline{c}'\underline{x}, \end{aligned} \tag{2.1}$$

where $\underline{x} \in R^n$.

A. Prékopa proved in [5],[6] that if the joint distribution of the random variables $\beta_1,\ldots,\beta_m$ is continuous and the joint density is of the following type

$$e^{-Q(\underline{z})}, \qquad \underline{z} \in R^m \tag{2.2}$$

where $Q(\underline{z})$ is a convex function in the entire space, which can be equal to $+\infty$, then the function $G(\underline{x})$ is logarithmic concave in the entire space R^n. In our model it is supposed that the joint distribution of the random variables $\beta_1,\ldots,\beta_m$ is a nondegenerated normal distribution. In this case their joint density is of the form

$$\frac{1}{(2\pi)^{\frac{m}{2}}\sqrt{|C|}}\, e^{-\frac{1}{2}\underline{z}'\, C^{-1}\underline{z}} \tag{2.3}$$

where the matrix C is the correlation matrix of the random variables $\beta_1,\ldots,\beta_m$, i.e.

$$c_{ik} = E(\beta_i\beta_k), \qquad i,k = 1,\ldots,m. \tag{2.4}$$

Since the distribution is non-degenerated, the matrix C is positive definite so C^{-1} exists and is also positive definite, hance by a well--known theorem the function

$$\underline{z}'\, C^{-1}\, \underline{z}, \tag{2.5}$$

is convex in the entire space R^m. If the joint density of the random variables $\beta_1,\ldots,\beta_m$ is (2.3), then the function $G(\underline{x})$ is logarithmic concave in the entire space R^n, i.e. for every pair of vectors $\underline{x}_1,\underline{x}_2 \in R^n$ and $0 < \lambda < 1$ we have the inequality

$$G(\lambda\underline{x}_1 + (1-\lambda)x_2) \geq [G(\underline{x}_1)]^{\lambda}[G(\underline{x}_2)]^{1-\lambda} \tag{2.6}$$

Since $G(\underline{x}) > 0$ for every $\underline{x} \in R^n$, this implies that log $G(\underline{x})$ is a finite valued concave function in the entire space R^n.

Algorithms for the solution of Problem (2.1). The function in the first constraint of (2.1) is quasi-concave. If a nonlinear programming method converges whenever the constraints are quasi-concave and the objective function is linear, it can be applied for the solution of our problem. Such a method is Zoutendijk's method of feasible directions in particular Procedure P2[10, p.74]. For the proof of the convergence in case of quasi-concave constraining functions and convex objective function (to be minimized) see [4], [5], [7]. Before presenting the method for the solution of Problem (2.1) we describe the mentioned Zoutendijk method for the problem

$$
\begin{aligned}
& G(\underline{x}) \geq p, \\
& \underline{a}_i' \underline{x} \geq b_i, \qquad i \in I, \\
& \min \ f(\underline{x}),
\end{aligned}
\tag{2.7}
$$

where the functions G and f are not specialized according to (1.1) and (1.5). We only suppose that these functions are differentiable in every variable in the entire space R^n. Suppose further, that the set determined by the linear constraints is nonempty and bounded. Let $\underline{x}_1$ be an arbitrary vector satisfying the constraints of Problem (2.7). We define by induction the successive iterations. Suppose we already determined the vectors $\underline{x}_1, \ldots, \underline{x}_k$ and want to determine the vector $\underline{x}_{k+1}$ in the (k+1) -th iteration. One iteration consists of two parts. In the first part we solve the following so called direction finding problem:

$$
\begin{aligned}
& G(\underline{x}_k) + \nabla G(\underline{x}_k)\,[\,\underline{x} - \underline{x}_k\,] + \upsilon y \geq p, \\
& \underline{a}_i' \underline{x} \geq b_i, \quad i \in I, \\
& \nabla f(\underline{x}_k)\,[\,\underline{x} - \underline{x}_k\,] \leq y, \\
& \min \ y,
\end{aligned}
\tag{2.8}
$$

where υ is an arbitrary but throughout the whole procedure fixed positive number. The number of the variables in the linear programming problem (2.8) is n+1, since the vector $\underline{x}$ has n components and y is a variable too. There is always at least one vector satisfying the constraints of (2.8). Such a vector is the n+1 component vector $\underline{x} = \underline{x}_k$, $y = 0$.

Since $\underline{x}$ varies in a bounded set, the objective function is bounded from below, hence the problem (2.8) has a finite optimum. Let y_{opt} denote the optimal value ot the objective function of Problem (2.8). If $y_{opt} = 0$, then the procedure terminates. If $y_{opt} \neq 0$, which means in this case $y_{opt} < 0$, then we proceed to the second part of the (k+1)-th iteration which is the determination of the steplength. Let $\underline{x}_k^*$ be an optimal solution of Problem (2.8). Then we minimize the function of the variable λ:

$$f(\underline{x}_k + \lambda(\underline{x}_k^* - \underline{x}_k)), \qquad \lambda \geq 0, \tag{2.9}$$

on that part of the halfline $\underline{x}_k + \lambda(\underline{x}_k^* - \underline{x}_k)$, which belongs to the set of feasible solutions of Problem (2.7) Under very general assumptions this minimum is obtained for some λ. If λ_k is a mininizing λ, then we define $\underline{x}_{k+1}$ as follows:

$$\underline{x}_{k+1} = \underline{x}_k + \lambda_k(\underline{x}_k^* - \underline{x}_k). \tag{2.10}$$

Solution of Problem (2.1) in two phases. In the second phase we solve Problem (2.1) under the assumption that we have a vector $\underline{x}_1$ satisfying the constraints. In the first phase our aim is to find a vector $\underline{x}_1$ satisfying the constraints.

In the second phase we take into account the meaning of the function G and the fact that f is a linear function of the form (1.5). The gradient of the function G will be given in the next section. The gradient of the function f is the constant vector $\underline{c}'$. The function (2.9) has the form

(2.11) $$\underline{c}' (\underline{x}_k + \lambda (\underline{x}_k^* - \underline{x}_k)).$$

This is to be minimized on a bounded interval since the set of vectors $\underline{x}$ defined by the constraints of Problem (2.1) is convex and bounded. Since we have $y_{opt} < 0$, it follows that $\underline{c}'\underline{x}_k^* < \underline{c}'\underline{x}_k$. Thus $\underline{x}_k^* = \underline{x}_k$ which implies that the set of vectors belonging to $0 \leq \lambda \leq 1$ is a non-degenerated interval. Since it is part of the above mentioned closed interval so the latter is also non-degenerated. The minimum is attained at that endpoint of the interval which belongs to the positive λ value. It is easy to see, that the vector $\underline{x}_{k+1}$ is a boundary point of the set defined by the constraints of the problem. Thus all the the points $\underline{x}_2, \underline{x}_3, \ldots$ are boundary points.

In the first phase we want to find a vector $\underline{x}_1$ satisfying the constraints of Problem (2.1). We use the same method which is described for the solution of Problem (2.7) but now we maximize the first constraining function G of Problem (2.1) under the linear constraints of Problem (2.1). This can be considered as a special case of Problem (2.7). Thus we apply the method of solution for the problem

(2.12) $$\underline{a}_i'\underline{x} \geq b_i, \qquad i \in I,$$
$$\max G(\underline{x}),$$

i.e. for the equivalent problem

(2.13) $$\underline{a}_i'\underline{x} \geq b_i, \qquad i \in I,$$
$$\min (-G(\underline{x}))$$

as long as we reach a vector $\underline{x}_1$ which satisfies the inequality

(2.14) $$G(\underline{x}_1) \geq p.$$

This vector $\underline{x}_1$ can be used as the starting vector for the second phase. The method applied for the problem (2.13) can be summarized the following manner. We start from a vector z_1 which satisfies the constraints

of the problem (2.13). Such a vector z_1 can easly be found by linear programming. If we already determined the vectors $z_1,\dots,z_k$ then in order to define z_{k+1} we consider the following direction finding problem:

$$\begin{aligned} &\underline{a}_i'\underline{z} \geq b_i, \qquad i \in I, \\ &-G(\underline{z}_k) + \nabla(-G)z_k)) \ (\underline{z}-\underline{z}_k) \leq y, \end{aligned} \tag{2.15}$$

which can be reformulated as follows

$$\begin{aligned} &\underline{a}_i' \ \underline{z} \ \geq b_i, \qquad i \in I, \\ &\min(-\nabla G(\underline{z}_k) \ (\underline{z} - \underline{z}_k)). \end{aligned} \tag{2.16}$$

The stopping rule and the step length determination was already discussed above in connection with Problem (2.7). After all the method applied in the first phase is the well-known gradient method.

3. Convergence of the procedure. As we already mentioned the convergence of the procedure applied for Problem (2.7) was considered in [4], [5], [7]. We recall the main theorem in a weaker but for the present purpose more comfortable form. This is the following

Theorem 1 Suppose that the following conditions hold:

I. The functions G and f are defined in the entire space R^n and have continuous gradient there.

II. The function G is quasi-concave and the function f is convex in the entire space.

III. The set $K = \{\underline{x} | \underline{a}_i'\underline{x} \geq b_i, \quad i \in I\}$ is non-empty and bounded.

IV. For every $\underline{x}$ satisfying the equality $G(\underline{x}) = p$ there correspondes a vector $\underline{y}$ in the set of feasible solutions

with the property

(3.1) $$\nabla G(\underline{x})\ (\underline{y} - \underline{x}) > 0.$$

If the procedure terminates in a finite number of steps and the last vector is $\underline{x}_N$, then we have

(3.2) $$f(\underline{x}_N) = \min_{\underline{x} \in L} f(\underline{x}).$$

If the sequence $\underline{x}_1, \underline{x}_2 \cdots$ is infinite then we have

(3.3) $$\lim_{K \to \infty} f(\underline{x}_k) = \min_{x \in L} f(\underline{x}),$$

where L is the set of the feasible solutions of the problem.

The convergence of the procedure applied for Problem (2.1) can be prooved by using Theorem 1. We formulate two different theorems for the two phases. It is obvious that the first phase has to terminate in a finite number of steps, while in the second phase it is enough if the procedure converges. First we consider the convergence problem of the second phase. We prove

Theorem 2. Besides the assumptions formulated in connection with Problem (2.1) suppose, that there exists a vector $\underline{y} \in L$, for which

(3.4) $$G(\underline{y}) > p.$$

Then the second phase procedure is either finite and the last vector $\underline{x}_N$ satisfies the relation (3.2) or it is infinite and the relation (3.3) holds. As before, L denotes the set of feasible solutions of the problem.

Proof. We have to prove that the assumptions I-IV of Theorem 1 hold. Assumptions I and II are satisfied trivially. Assumption III was earlier introduced in connection with Problem (*2.1*). So we only have to check the validity of Assumption IV. Indirect proof will be applied. Suppose there exists an $\underline{x} \in L$ for which $G(\underline{x}) = p$ such that for every $\underline{y} \in D$ we have

$$\nabla G(\underline{x}) \ (\underline{y} - \underline{x}) \leqq 0. \tag{3.5}$$

The function log G is finite valued and concave in the entire space. This implies that the following inequality holds for every $\underline{y} \in R^n$

$$\log G(\underline{y}) - \log G(\underline{x}) \leqq \log G(\underline{x}) \ (\underline{y} - \underline{x}) = \tag{3.6}$$
$$= \frac{1}{G(\underline{x})} \nabla G(\underline{x}) \ (\underline{y} - \underline{x}).$$

If $\underline{y} \in L$, then (*3.5*) and the inequality $G(\underline{x}) > 0$ together imply

$$G(\underline{y}) \leqq G(\underline{x}) = p. \tag{3.7}$$

This is a contradiction, thus the theorem is proved.

The following theorem summarizes our statement in connection with the finiteness of the first phase.

Theorem 3. Besides the assumptions already introduced in connection with Problem (*2.1*) assume that there exists a vector $\underline{y} \in K$, for which the inequality (*3.4*) holds. Then starting from any vector $\underline{z}_1 \in K$, after a finite number of steps we reach a vector lying in the set L.

Proof. The method applied in the first phase is the classical gradient method, so we might refer to any already existing convergence theorem relative to the gradient method. Nevertheless, for the sake of presenting a unified approach we refer to Theorem 1. By Theorem 1 it follows that if we apply the method of the first phase not for Problem (*2.13*) but for the problem

$$(3.8) \qquad \begin{aligned} & \underline{a}_i' \underline{z} \geq b_i, \qquad i \in I, \\ & \min\ (-\log G(z)), \end{aligned}$$

then the obtained sequence $\underline{z}_1, \underline{z}_2, \ldots$ is either finite and the last vector $\underline{z}$ minimizes the objective function of Problem (*3.8*) or the following relation holds

$$(3.9) \qquad \lim_{k \to \infty} (-\log G(\underline{z}_k)) = \min_{\underline{z} \in K} (-\log G(z)).$$

Now the (*2.16*) - type problem is the following

$$(3.10) \qquad \begin{aligned} & \underline{a}_i' \underline{z} \geq b_i, \qquad i \in I, \\ & \min\ \left(-\frac{1}{G(\underline{z}_k)} \nabla G(\underline{z}_k)\ (\underline{z} - \underline{z}_k)\right). \end{aligned}$$

The objective functions of Problems (*3.8*) and (*2.16*) differ only in a positive constant factor hence the sets of optimal solutions are the same. The same is the situation in the second part of the k-th iteration when we determine the step length as it makes no difference whether the function $-G$ or the function $-\log G$ is minimized. Thus if we obtain a sequence $\underline{z}_1, \underline{z}_2, \ldots$ by the procedure applied for Problem (*2.13*), this sequence is appropriate from the point of view of the procedure applied for Problem (*3.8*) too hence in case of a finite sequence the last vector minimizes the function $-\log G$, so at the same time the function $-G$ and in case of an infinite sequence

(3.11) $$\lim_{k\to\infty} \left(-G\left(\underline{z}_k\right)\right) = \min_{\underline{z}\in K} \left(-G(\underline{z})\right).$$

Since there exists a $\underline{y}\in K$, for which $G(\underline{y}) > p$ holds after a finite number of steps we must reach a vector lying in the set L. Thus the theorem is proved.

4. Evaluation of the gradient of the nonlinear constraining function. Both in the first and the second phase of the procedure applied for the solution of Problem (2.1) we need values of the gradient of the function $G(\underline{x})$. In addition we need the values of the function $G(\underline{x})$. To the problem of determination of these latter values we return Section 6. Now we show that the method we can apply for the determination of $G(\underline{x})$ is essentially suitable for the determination of $\nabla G(\underline{x})$. We shall see that while the evaluation of $G(\underline{x})$ requires the evaluation of the distribution function of the m dimensional normal distribution the determination of $\nabla G(\underline{x})$ requires the evaluation of the distribution function of the m-1 dimensional normal distribution.

Denote $\phi(\underline{z};C)$ the probability density function (2.3) and $\underline{\phi}(\underline{z};C)$ the corresponding probability distribution function. Introduce the following notations

$$L_i(\underline{x}) = \frac{1}{\sigma_i}\left(\underline{a}_i^{\prime}\underline{x} - b_i\right), \qquad i = 1,\ldots,m,$$

(4.1) $$\underline{L}(\underline{x}) = \begin{pmatrix} L_1(\underline{x}) \\ \vdots \\ L_m(\underline{x}) \end{pmatrix}.$$

The function G can be written now as

(4.2) $$G(\underline{x}) = \phi\left(\underline{L}(\underline{x}); C\right).$$

It is well-known in probability theory that if $F(\underline{z}) = F(z_1,\ldots,z_m)$ is a distribution function of a continuous distribution, which is the joint distribution function of the random variables $\xi_1,\ldots,\xi_m$ then between the conditional distribution function $F(z_2,\ldots,z_m \mid z_1)$ of the random variables $\xi_2,\ldots,\xi_m$ given that $\xi_1 = z_1$ and the partial derivative of the function F with respect to z_1 the following relation holds

$$\frac{\partial F(z_1,\ldots,z_m)}{\partial z_1} = F(z_2,\ldots,z_m|z_1)\ f_1(z_1), \tag{4.3}$$

where $f_1(z)$ is the density function of the random variable ξ_1. Similar equalities hold for the derivatives with respect to the other variables. Applying the formula (4.3) for the distribution function $\phi(\underline{z};C)$ we obtain

$$\frac{\partial \phi(\underline{z};C)}{\partial z_1} = \phi(z_2,\ldots,z_m|z_1)\phi(z_1) \tag{4.4}$$

where $\rho(z)$ is the density function of the standard normal distribution. It is also well-known, that

$$\phi(z_2,\ldots,z_m|z_1)=\phi\left(\frac{z_2-r_{12}z_1}{\sqrt{1-r_{12}^2}},\ldots,\frac{z_m-r_{1m}z_1}{\sqrt{1-r_{1m}^2}}\ ;\ S^{(1)}\right), \tag{4.5}$$

where the correlation matrix $S^{(1)}$ consists of the following elements

$$s_{ik}^{(1)} = \frac{r_{ik} - r_{i1}\,r_{k1}}{\sqrt{1 - r_{i1}^2}\ \sqrt{1 - r_{k1}^2}} \qquad i,k=2,\ldots,m. \tag{4.6}$$

Similar formulas hold if one of the variables $z_2,\ldots,z_m$ plays the role of z_1. The corresponding correlation matrices will be denoted by $S^{(2)},\ldots,S^{(m)}$.

It is reasonable to introduce a notation for the components of the vectors $\underline{a}_1,\ldots,\underline{a}_m$. Let us denote the components of $\underline{a}_j$ by the symbols $a_{j1},\ldots,a_{jn}$, $j=1,\ldots,m$. Now we give a formula for $\nabla G(\underline{x})$. In order to avoid inconvenient notation, we describe the components of this vector, but do not unite them in a vector form. These components are the following

$$\phi\left(\frac{L_2(\underline{x})-r_{12}L_1(\underline{x})}{\sqrt{1-r_{12}^2}};\ldots,\frac{L_m(\underline{x})-r_{1m}L_1(\underline{x})}{\sqrt{1-r_{1m}^2}};S^{(1)}\right)\phi(L_1(\underline{x}))\frac{a_{11}}{\sigma_1}+\ldots+$$

$$+\phi\left(\frac{L_1(\underline{x})-r_{m1}L_m(\underline{x})}{\sqrt{1-r_{m1}^2}},\ldots,\frac{L_{m-1}(\underline{x})-r_{m\,m-1}L_m(\underline{x})}{\sqrt{1-r_{m\,m-1}^2}};S^{(m)}\right)\phi(L_m(\underline{x}))\frac{a_{m1}}{\sigma_m},$$

(4.7)

$$\vdots$$

$$\phi\left(\frac{L_2(\underline{x})-r_{12}L_1(\underline{x})}{\sqrt{1-r_{12}^2}},\ldots,\frac{L_m(\underline{x})-r_{1m}L_1(\underline{x})}{\sqrt{1-r_{1m}^2}};S^{(1)}\right)\phi(L_1(\underline{x}))\frac{a_{1n}}{\sigma_1}+\ldots+$$

$$+\phi\left(\frac{L_1(\underline{x})-r_{m1}L_m(\underline{x})}{\sqrt{1-r_{m1}^2}},\ldots,\frac{L_{m-1}(\underline{x})-r_{m\,m-1}L_m(\underline{x})}{\sqrt{1-r_{m\,m-1}^2}};S^{(m)}\right)\phi(L_m(\underline{x}))\frac{a_{mn}}{\sigma_m}.$$

The numerical determination of the values of the function ϕ does not make any difficulty. After all the same method applies for the determination of $G(\underline{x})$ and $\nabla G(\underline{x})$.

5. Formulation of the economic problem. At the Institute of Economic Planning belonging to the Hungarian National Planning Office a large scale linear programming model was developed as a part of the planning method of the 4th Five-Year Plan. The model is a large-scale linear programming model comprising the interrelations of the physical values and the financial processes. It has a decomposition structure i.e. the model contains sectors related to the branches of the national economy. Our underlying deterministic model is the electrical energy sector model of the large-scale model. When formulating our model the remaining sectors were supposed to work on fixed levels. Specialities of the electrical energy sector were taken into account such as the long time-lag of energetical investments, the substitution possibility of different kinds of fuels etc. The variables of the model can be classified in the following manner: production of electrical energy by exogenous capacity (having completed before the planning period), production of electrical energy by endogenous capacity (to be completed during the planning period), the various kinds of fuels, import and export of electrical energy, marking out the Rouble and Dollar relations, variables representing the individual investment projects (new power stations), investment variables which take into account the different financial resources, purposes and manners of use. The constraints of the model contain the manpower balances, constraints on investment, foreign trade balances, the balance of the state budget, the constraints containing the demand for electrical energy and other financial constraints.

The data of the numerical model were gained from the above mentoined large scale model, from the coordination period of planning works and from the official statistics. Since the plan indices are obtained by forecasting, we have to deal with uncertainty so the formulation of a stochastic model is reasonable. We assumed that the uncertainty in the statistical data can be neglected as compared to that of the plan indices. Having analysed their economic contents, four constraints of the deterministic model was regarded stochastic, i.e. $m = 4$ in the model (2.1).

Now we outline the content of the mentioned four constraints and the four right hand side values in the deterministic model. The right hand side value b_1 is the planned negative saldo of the foreign trade in

Rouble relation while the corresponding constraint prescribes that this saldo should not exceed a certain planned level. The right hand side value b_2 and the corresponding constraint have the same meaning in Dollar relation. The underlying deterministic version of the third and fourth stochastic constraints express the relation between the electrical energy sector and the other sectors. The third constraint is essentially that row of the input-output table which corresponds to the electrical energy sector. This constraint includes the requirement for the electrical energy sector to assure the equilibrium of the input- and output of the national gross product in value terms. b_3 is equal to the sum of inputs of the productive sectors except for the electrical energy sector and of the non-productive sectors effected by the electrical energy sector. Finally the right hand side value b_4 is equal to the cummulative minimum demand for electrical energy in the productive sectors except for the electrical energy sector and the non-productive sectors. The corresponding constraint is the product balance of the electrical energy in natural units. The right hand side values of the stochastic constraints in the stochastic programming model are $b_1 + \sigma_1\beta_1$, $b_2 + \sigma_2\beta_2$, $b_3 + \sigma_3\beta_3$, $b_4 + \sigma_4\beta_4$. The joint distribution of these random variables was supposed to be normal. Its parameters are given in Section 7.

6. Brief description of the computer program. The detailed description of the computer program of Model (*2.1*) can be found in [2]. Here we outline only the most important features.

The main problem was the numerical determination of the values of the joint distribution function of the random variables $\beta_1,\ldots,\beta_m$. We recall, that the determination of these values and of the gradient values is essentially the same problem (see (*4.2*) and (*4.7*)). The determination of the values of the function ϕ is carried out by a subroutine using a modified Monte Carlo integration technique. We approximate the integral of the function $\phi(\underline{z};C)$ over the set $\{\underline{z} \mid \underline{z} \leq \underline{u}\}$ in the following way. First we choose a vector $\underline{u}_o$ so that the integral over the set $\{\underline{z} \mid \underline{u}_o \leq \underline{z} \leq \underline{u}\}$ should be very near the integral over the former set. Then we choose uniformly distributed random points from the set $\{\underline{z} \mid \underline{u}_o \leq \underline{z} \leq \underline{u}\}$ and take the arithmetic mean of the function values

belonging to these points. It is possible to determine the number of random points resulting a required precision i.e. incase of which the relative error does not exceed a prescribed level. In the special model we have to deal with the four dimensional normal distribution. The computation of one value of the distribution function takes 0.5 sec if the prescribed upper bound for the relative error 5%.

Special attention should be paid to the determination of the step length at each iteration. Our procedure applied for the evaluation of the distribution function of the normal distribution (like any other available procedure) does not give exact results. Only the expectation of the estimations are equal to the true values. This fact is very inconvenient when determining the feasible part of the ray $\{\underline{x}_k + \lambda(\underline{x}_k^* - \underline{x}_k) \mid \lambda \geq 0\}$ i.e. the intersection of this ray with the boundary of the set of feasible solutions. We apply here a stepping procedure moving up and down along the ray and successively reduce these length between the scaling points. In this way we are lead near the required intersection point.

We had to agree in an "optimality criterion". A criterion formulated only in terms of the value y_{opt} may have been satisfactory. However, we thought to be more suitable to formulate a more rigorous stopping rule, namely the following. The vector $\underline{x}_k$ is considered optimal if the difference between the values of the objective function belonging to $\underline{x}_{k+1}$ and $\underline{x}_k$ does not exceed 1% of the latter and at the same time the difference between the individual components does not exceed 2% of the components of $\underline{x}_k$.

The computations were executed on a CDC 3300 computer of the Hungarian Academy of Sciences. The program consists of a main program and six overlays. Five overlays run the simplex method, the sixth one contains the algorithm described in Section 2 and the evaluation procedure of the values of the normal distribution function.

The problem run in two very similar programs, one of them executed the first phase and the other one executed the second phase.

7. Numerical data and results. The special model is of the following form:

$$G(\underline{x}) = P(\underline{a}_i' \underline{x} \leq \sigma_i \beta_i + b_i, \quad i = 1,2,3,4) \leq p$$

(7.1)
$$\underline{a}_i' \underline{x} \geq b_i, \quad i = 5,\ldots,110,$$
$$\min \underline{c}' \underline{x}.$$

The linear functions standing on the left hand sides of the stochastic constraints are specialized as follows:

$$\underline{a}_1' \underline{x} = -25x_{25},$$
$$\underline{a}_2' \underline{x} = -16.67x_{26},$$
$$\underline{a}_3' \underline{x} = 0.8696x_{24} + x_{40},$$
$$\underline{a}_4 \underline{x} = 0.9(x_1 + x_2 + x_6 + x_7) - 0.115x_{24}$$

where x_1 and x_2 are the productions of the electrical energy by exogenous resp. endogenous capacity, x_6 and x_7 are the imports of electrical energy in Rouble resp. Dollar relation, x_{24} is the value of the production of the electrical energy sector, x_{25} and x_{26} are the values of the imports of the electrical energy sector in Rouble resp. Dollar relation, x_{40} is the total value of the import of the electrical energy sector.

The expectations b_1, b_2, b_3, b_4 and the dispersions $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ on the right hand sides of the stochastic constraints are the following:

(7.3)
$$b_1 = -48313, \quad \sigma_1 = 483,$$
$$b_2 = -426, \quad \sigma_2 = 4,$$
$$b_3 = 16000, \quad \sigma_3 = 160,$$
$$b_4 = 19000, \quad \sigma_4 = 195.$$

The expectations of the random variables $\beta_1, \beta_2, \beta_3, \beta_4$ are equal to zero,

their dispersions are equal to 1 and their correlation matrix is the following matrix C:

(7.4)
$$C = \begin{pmatrix} 1 & -0.8 & 0.4 & 0.4 \\ -0.8 & 1 & 0.1 & 0.1 \\ 0.4 & 0.1 & 1 & 0.9 \\ 0.4 & 0.1 & 0.9 & 1 \end{pmatrix}.$$

The linear functions in the second row of Problem (7.1) are divided into two groups. The first group contains the linear constraints having subscripts i=5,...,52, the second group contains the remaining linear constraints of subscripts i=53,...,110. These latters are lower and upper bounds for some variables and nonnegativity constraints for the other variables. These and further numerical data of Problem (7.1) can be found in the paper [9], namely under (7.5) of [9], and in the appendices [3] and [4] of [9].

The objective function which is the profit multiplied by -1, is the following

(7.5)
$$\underline{c}'\underline{x} = x_{35} - x_{36},$$

where x_{35} is the increase of wages and x_{36} is the enterprise profit before taxation. The prescribed probability level is p=0.9.

First phase. We used $\underline{x}_{lin}$, the optimal solution of the deterministic underlying problem as the initial feasible solution. We computed the value of the function G belonging to this vector and obtained

(7.6)
$$G(\underline{x}_{lin}) = 0.09.$$

So the optimal solution $\underline{x}_{lin}$ guarantees only a very low reliability level for the system. After this we maximized the function $G(\underline{x})$ under the linear constraints of Problem (7.1). Five iterations were performed. The following numbers were obtained as values of the function $G(\underline{x})$ (the

first belongs to the vector $\underline{x}_{lin}$):

0.09; 0.13; 0.72; 0.90; 0.94; 0,97.

The first phase may have been terminated when we first received a probability greater than 0.9. Thus four iterations would have been enough. However, we were interested in answering the question that how high probability level could be achieved under the linear constraints of Problem (7.1). We interrupted the computations at the value 0.97 because it already shew that this maximal probability level is very high. The first phase was executed in 19 minutes, namely the compute time was 8 min 49 sec, and the channel time was 5 min 19. sec.

Second phase. As starting vector we used the last obtained vector in the first phase in case of which the value of the function G equals 0.97. The computer program run 46 minutes out of which the compute time was 25 min and the channel time was 12 min.

Deatailed description of the 46 minutes total running time.

Preparing the data for the simplex methods	3 min	10 sec
Running of the simplex methods	30 min	39 sec
Running of the stepping procedures	6 min	34 sec
Checking the optimality criteria and other computations	2 min.	

The optimization was performed in 9 steps. The values of the objective function in the iterations were the following.

- 4033; -4101; -4366.9; -4367; -4367,32;
- 4367,48;-4367.84;-4367,9; -4369.71; -4369,86.

It is surprising that on the optimal solution $\underline{x}_{stoch}$ of the stochastic programming model the value of the objective function is equal to the value of the objective function on $\underline{x}_{lin}$. As for the values of the function G we have $G(\underline{x}_{stoch}) = 0.9$ and we recall that $G(\underline{x}_{lin}) = 0.09$.

Thus it is possible to achieve the same objective by a vector representing a considerably greater reliability level than $\underline{x}_{lin}$. This phenomenon is remarkable apart from the special meaning of the numerical model.

Below we give those components of the vectors $\underline{x}_{lin}$ and $\underline{x}_{stoch}$ the relative differences of which are more than 10%:

Component subscript	x_{lin}	$\underline{x}_{stoch} = \underline{x}^{(1)}_{stoch}$
20	0	1233,9
21	994	13,7
22	1950	714,4
23	517	1586,2
43	2370	1655,8
46	2407	1007,3

The economic meanings of these components are described at the end of the paper.

The complete computer program run with two more data-set as an experiment. Here under a.) and b.) we make an account about the obtained results.

a) Except for the probability p the data remained unchanged. The new value was somewhat greater, namely p = 0.95. The value of the objective function is -4365.8 on the optimal solution of the stochastic programming model. Thus there is no great difference relative to the former optimal value of the objective function. The optimal solution differs considerably from both the former vector $\underline{x}_{lin}$ and the former vector $\underline{x}^{(1)}_{stoch}$. Here we list those components in which the relative deviation is at least 10%.

Component subscript	$\underline{x}_{lin}$	$\underline{x}^{(2)}_{stoch}$ (the case of p=0.95)
5	0	0.353
10	0.1	0.003
13	0.37	0.44
20	0	11.1
21	994	1080
41	830	938
46	2407	1.24

b) In this case the reliability level is again $p = 0.9$ and except for the correlation matrix the data remain unchanged. We have the following new correlation matrix.

$$C_1 = \begin{pmatrix} 1 & -0.7 & 0.3 & 0.3 \\ -0.7 & 1 & 0.1 & 0.1 \\ 0.3 & 0.1 & 1 & 0.9 \\ 0.3 & 0.1 & 0.9 & 1 \end{pmatrix} . \tag{7.7}$$

The optimal value of the objective function is -4292 which considerably differs from the former optimal values of the objective function. The components in which the relative deviation exceeds 10% are the following.

Component subscript	$\underline{x}_{lin}$	$\underline{x}^{(3)}_{stoch}$ (the case of the matrix C_1)
5	0	10.1
12	0.23	0.11
13	0.37	0.5
20	0	341
21	994	879
22	1950	1602
23	514	720
41	830	930
45	2962	2429
46	2407	38

The detailed economic analysis of the vectors $\underline{x}_{lin}$, $\underline{x}^{(1)}_{stoch}$ $\underline{x}^{(2)}_{stoch}$ $\underline{x}^{(3)}_{stoch}$ is not our aim in this paper. The only thing what we mention is that the plans of greater reliability level propose the use of more coal and less fuel oil to be consumed in endogenous power stations.

Finally we describe the meanings of those components of the optimal solutions which are involved in this section.

Component subscript	Meaning of the component
5	Production of electrical energy by endogenous capacity in nuclear power station.
10	Individual investment project "Dunamenti II" to be completed in 1977.
12	Individual investment project of a nuclear power station at Paks, to be completed in 1978.
13	Individual investment project of an oil operated power station, to be completed in 1978.
20	Consumption of natural gas in exogenous power station.
21	Consumption of fuel oil in exogenous power station.
22	Consumption of natural gas in endogenous power station.
23	Consumption of fuel oil in endogenous power station.
41	Total machinery requirement of investments in the electrical energy sector.

43	Investment surplus in 1972
45	Investment surplus in 1974
46	Investment surplus in 1975.

REFERENCES

[1] G.Báger, T.Morva, L.Szabó. A model of the 4th Five-Year Plan, which comprises the natural, value and financial processes,
Országos Tervhivatal, Tervgazdasági Intézetének Közleményei 6 (1971) 55-130 (in Hungarian).

[2] I. Deák. Computer evaluation of a stochastic programming model, MTA Számitástechnikai Központ Közleményei 9 (1972) 33-49 (in Hungarian).

[3] S. Ganczer (ed). National economic planning and programming, Publishing House of Economics and Law, Budapest, (1973) (in Hungarian).

[4] A. Prékopa. On probabilistic constrained programming, in Proceedings of the Princeton Symposium on Mathematical Programming (Princeton University Press, Princeton,N. 1970) 113-138.

[5] A. Prékopa. On optimization problems of stochastic systems, Hungarian Academy of Sciences, Budapest, 1970 (in Hungarian).

[6] A. Prékopa. Logarithmic concave measures with applications to stochastic programming, Acta Scientiarum Mathematicarum 32 (1971) 301-316.

[7] Extension of the method of feasible directions for quasi-concave constraining functions, **MTA Számitástechnikai Központ Közleményei 9 (1972), 3-16** (in Hungarian), Mathematische Operationsforschung und Statistik 5 (1974) 281-293 (in German).

[8] A. Prékopa: Contributions to the theory of stochastic programming, Mathematical Programming 4 (1973) 202-221.

[9] A. Prékopa, S.Ganczer, I.Deák, K.Patyi. The STABIL stochastic programming model and its experimental application to the electrical energy sector of the Hungarian economy, **Alkalmazott Matematikai Lapok** 1(1975) 3 - 22 (in Hungarian).

[10] G. Zoutendijk. Methods of feasible directions Elsevier Publishing Company, Amsterdam and New York, 1960.

A Nonlinear Programming Method for the Solution of a Stochastic Programming Model of A. Prékopa

by

J. Mayer

COMPUTER AND AUTOMATION INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES

1. Introduction

The nonlinear programming algorithm presented in this paper is designed for the solution of a stochastic programming problem of A. Prékopa [6], [8],[10], but it can also be applied for nonlinear programming problems with only one nonlinear constraint. It is a combination of Zoutendijk's P1 method [14] and the reduced gradient method of Wolfe [13], Abadie and Carpentier [1]. The algorithm works using feasible directions, but for handling the linear constraints, the very effective techniques of the reduced gradient method are applied.

In the next section the stochastic programming problem of A. Prékopa is discussed from a nonlinear programming point of view. Section 3 contains some numerical experiences with GRG and the motivation of our method. In the subsequent section the algorithms is described in details, and its convergence investigated. The last section contains some numerical experiences with the method.

2. The stochastic programming model of A. Prékopa

We consider the question of solving the following stochastic programming model of A. Prékopa:

$$\begin{aligned} \max \quad & f(\underline{x}) \\ & G(\underline{x}) \geq p \\ & A\,\underline{x} = \underline{b} \\ & \underline{x} \geq \underline{0}. \end{aligned} \tag{1}$$

where $\underline{x} \in R^n$, $\underline{b} \in R^m$, A is $m \times n$, and $m < n$, $r(A)=m$, matrix A is of full rank. The objective function is assumed to be a concave, continuously differentiable function and the nonlinear constraint has the following form:

$$G(\underline{x}) = P\{S\,\underline{x} \geqq \underline{\beta}\} \text{ , where } S \text{ is an } r \times n \text{ matrix,}$$

$\underline{\beta} = (\beta_1, \ldots, \beta_r)$ is a random vector-variable, i.e. $G(\underline{x})$ is the probability of the event, that the system of linear inequalities $S\underline{x} \geqq \underline{\beta}$ with random right-hand sides is fulfilled.

From the nonlinear programming point of view (1) is a nonlinear programming problem of a special type, it contains only one nonlinear constraint. The nonlinear programming approaches to the solution of the problem are all based on the following fundamental result of A. Prékopa: [6] , [7] , [8],

Theorem 1. Let $\beta_1, \ldots, \beta_r$ be random variables with a joint probability density function of the following form: $h(\underline{z}) = e^{-Q(\underline{z})}$, where $\underline{z} \in R^r$, $Q(\underline{z})$ is a convex function in the entire space R^r. Then the function $G(\underline{x})$ in (1) is a logarithmic concave function on the entire space R^n.

In this paper it is assumed throughout that $\beta_1, \ldots, \beta_r$ are in general not independent random variables with a normal joint probability density function. It follows at once, that $G(\underline{x})$ has the following attractive properties:

Lemma 1. Under the above assumptions:

a.) $G(\underline{x})$ is a logarithmic concave function.

b.) $\nabla G(\underline{x})$ is Lipschitz-continuous, i.e. there exists a constant $L > 0$, such that for every $\underline{x} \in R^n$, $y \in R^m$

$$|| \nabla G(\underline{x}) - \nabla G(\underline{y}) || \leq L \cdot || \underline{x} - \underline{y} || \text{ holds.}$$

c.) If there exists a feasible $\underline{y}$ such that $G(\underline{y}) > p$ then for problem (1) the Kuhn-Tucker conditions are necessary and

sufficient conditions of optimality.

Proof: Statement a.) is a direct consequence of Theorem 1, and b.) follows easily from the nice properties of the normal probability distribution function. For c.) Prékopa proved [9], that the existence of a vector $\underline{y} \in R^n$, for which $G(\underline{y}) > p$ holds implies that the Kuhn-Tucker constraint qualification holds on the whole set of feasible solutions of problem (1). From this fact it is easy to derive c.) using the classical theory of optimality conditions.

3. Motivation of the algorithm

In this section we summarize our computational experiences with GRG reported in [5], and give some motivation of the algorithm presented in the next section.

There are two main difficulties which one encounters trying to implement some nonlinear programming technique to solve problem (1).

a.) The evaluation of $G(\underline{x})$ requires the calculation of the r-dimensional normal probability distribution function, i.e. numerical integration in r-dimensional space. From the numerical study made by Deák [2] it turns out, that the best way to compute $G(\underline{x})$ is using simulation techniques. We also used his subroutine for the calculation of the stochastic constraint in our numerical experimentation.

b.) The nonlinear programming problem (1) contains a constraint, which can be calculated only approximately, and this calculation is relatively time-consuming. The purpose of this paper is to discuss this second difficulty.

Our first attempt to solve (1) was using a general purpose experimental GRG code, a failure. Namely at every iteration GRG first generates a point, which is in general not feasible, then it tries to return to the feasible surface by initiating a procedure based on Newton's method. However in the

necessary matrix inversion computing errors get out of control, and afterwards it is impossible to return to the feasible surface.

Next we developed a specialized version of GRG, which is based on the product form inverse-base technique for the linear constraints, and the critical step, i.e. returning to the feasible surface, reduces to the problem of computing the intersection of a line, completely determined by the linear constraints, and the surface $G(\underline{x}) = p$. This is a much more stable approach, but the algorithm requires usually too much evaluations of the stochastic constraint in locating the surface. As the calculation of $G(\underline{x})$ is time-consuming, the code uses too much computing time.

Guided by this experiences we decided to modify the reduced gradient method. In the new method only feasible points are generated, and the efficient way of handling linear constraints, inherent in the reduced gradient method is preserved. To explain the idea, let us assume that at some point $\underline{x}$, if we take the partition of the variables $\underline{x} = (\underline{y},\underline{z})$, $\underline{y} \in R^m$, and the same partition for A as $A = (B,C)$ where B is $m \times m$, then B is nonsingular and $\underline{y} > \underline{0}$.

We consider two different ways for determination a direction of move at $\underline{x}$. Let us denote the direction by $\underline{w}$, and take the same partition as for $\underline{x}$: $\underline{w} = (\underline{u},\underline{v})$, with $\underline{u} \in R^m$.

The corresponding subproblem in Zoutendijk's P1 method:

$$\max \xi$$

$$\nabla_{\underline{y}}^T f(\underline{x})\underline{u} + \nabla_{\underline{z}}^T f(\underline{x})\ \underline{v} \geq \xi$$

$$\nabla_{\underline{y}}^T G(\underline{x})\underline{u} + \nabla_{\underline{z}}^T G(\underline{x})\ \underline{v} \geq \nu\xi, \qquad \text{if} \quad G(\underline{x}) = p$$

$$B\underline{u} + C\underline{v} = \underline{0}$$

$$v_i \geq 0 \qquad \text{if} \quad z_i = 0, \quad i = 1,2,\ldots,n-m$$

$$||(\underline{u},\underline{v})|| \leq 1$$

where $\nu > 0$ is a fixed number during the procedure.
The subproblem in GRG:

$$\max \xi$$

$$\nabla_{\underline{y}}^T f(\underline{x})\ \underline{u} + \nabla_{\underline{z}}^T f(\underline{x})\ \underline{v} \geq \xi$$

$$\nabla_{\underline{y}}^T G(\underline{x})\ \underline{u} + \nabla_{\underline{z}}^T G(\underline{x})\ \underline{v} \geq 0 \qquad \text{if} \quad G(\underline{x}) = p$$

$$B\underline{u} + C\ \underline{v} = \underline{0}$$

$$v_i \geq 0 \qquad \text{if} \quad z_i = 0, \quad i = 1,\ldots,n-m$$

$$||\ \underline{v}\ || \leq 1$$

In the new method the following subproblem is used determining the direction of move:

$$\max \xi$$

$$\nabla_{\underline{y}}^T f(\underline{x})\ \underline{u} + \nabla_{\underline{z}}^T f(\underline{x})\ \underline{v} \geq \xi$$

$$\nabla_{\underline{y}}^T G(\underline{x})\ \underline{u} + \nabla_{\underline{z}}^T G(\underline{x})\ \underline{v} \geq \nu\xi \qquad \text{if} \quad G(\underline{x}) = p$$

$$B\underline{u} + C\underline{v} = \underline{0}$$

$$v_i \geq 0 \qquad \text{if} \quad z_i = 0, \quad i = 1,\ldots,n-m$$

$$||\underline{v}|| \leq 1$$

This form enables a similar reduction like in the reduced gradient method, and in addition the algorithm will generate feasible points. So the linear part is handled by storing and updating B^{-1}, performing a change of the base if necessary, while feasibility is maintained. The algorithm constructed this way would be subject to jamming, so an anti zig-zagging prevention must be incorporated.

4. Algorithm for the solution of the stochastic programming problem

In this section we describe a nonlinear programming method for the solution of problem (1). It is also applicable for nonlinear programming problems having only one nonlinear constraint, provided that the objective function and constraints fulfil the conditions listed below.

The convergence of the algorithm can certainly be proved under less restrictive assumptions, these assumptions were selected because in our practical problems they were fulfilled and on the other hand they allow a short and elegant convergence proof. For the sake of completeness our list also contains some already verfied properties of $G(\underline{x})$. The assumptions are as follows:

a.) The objective function $f(\underline{x})$ is concave, differentiable, with a Lipschitz-continuous gradient, i.e. there exists a constant $L > 0$ such that for every $\underline{x} \in R^n$, $\underline{y} \in R^n$

$$||\nabla f(\underline{x}) - \nabla f(\underline{y})|| \leq L \cdot ||\underline{y} - \underline{x}|| \qquad \text{is fulfilled.}$$

b.) $G(\underline{x})$ is a logarithmic concave differentiable function, $\nabla G(\underline{x})$ is Lipschitz-continuous and bounded on R^n.

c.) There exists a feasible point $\underline{y}$ such that $G(\underline{y}) > p$ holds.

d.) The set of feasible solutions of problem (1) is bounded.

e.) At any point $\underline{x}$ of the feasible domain there exists a base B of A such that for all $i \in I_B$, $x_i > 0$ holds, where I_B is the set of basic indices. In other words we exclude degeneracy. Assume further, that $r(A) = m$, matrix A is of full rank.

The algorithm starts from a feasible point $\underline{x}^{(1)}$, furthermore $\varepsilon_1 > 0$ and the index set I_1 are selected in such a way that the corresponding columns of A form a base, and for all $i \in I_1$, $x_i^{(1)} > \varepsilon_1$ is fulfilled. Let $\nu > 0$, $R > 0$ be fixed numbers during the procedure.

The algorithm works in an iterative manner, at each iteration new feasible solution $\underline{x}^{(k)}$, new tolerance ε_k and new index set I_k are determined, k=1,2,..., and each iteration is divided into steps. For the sake of simplicity the base corresponding to I_k is denoted by B, the linear part of the constraints is handled by storing and updating B^{-1}.

Let us assume that $\underline{x}^{(k)}, \varepsilon, I_k$, having the same properties as for k=1, are determined, $k \geq 1$. For notational simplicity assume that the base consists of the first m colums of A, i.e. $I_k=\{1,\ldots m\}$, denote the base by B. All vectors and the matrix A are partitioned into basic and nonbasic part: $\underline{x}^{(k)} = (\underline{y}^{(k)}, \underline{z}^{(k)})$, $A=(B,C)$, similarly for the direction $\underline{w}^{(k)}=(\underline{u}^{(k)}, \underline{v}^{(k)})$ where $\underline{y}^{(k)} \in R^m$, $\underline{u}^{(k)} \in R^m$ and B is $m \times m$.

The next iteration consits of the following steps:

Step 1: Compute the reduced gradient of $f(\underline{x})$ and $G(\underline{x})$ with respect to the linear equality constraints:

$$
\begin{aligned}
\underline{r}^T &= \nabla_{\underline{z}}^T f(\underline{x}^{(k)}) - \nabla_{\underline{y}}^T f(\underline{x}^{(k)}) \quad B^{-1}C \\
\underline{s}^T &= \nabla_{\underline{z}}^T G(\underline{x}^{(k)}) - \nabla_{\underline{y}}^T G(\underline{x}^{(k)}) \; B^{-1}C
\end{aligned} \tag{2}
$$

Step 2: Solve the following direction-finding subproblem:

$$
\begin{aligned}
&\max \xi \\
&\underline{r}^T \underline{v} \geq \xi \\
&\underline{s}^T \underline{v} \geq \nu\xi \qquad \text{if} \quad G(\underline{x}^{(k)}) \leq p + \varepsilon_k \\
&v_i \geq 0 \qquad \text{if} \quad z_i^{(k)} \leq \varepsilon_k, \quad k=1,\ldots,n-m \\
&||\underline{v}|| \leq 1
\end{aligned} \tag{3}
$$

Let the optimal sulution of this problem be denoted by $(\hat{\underline{v}}, \hat{\xi})$.

Step 3: If $\hat{\xi} > \varepsilon_k$ continue at Step 4. Otherwise we consider two cases:

Case 1: If $\hat{\xi} \leq \varepsilon_k$ and $\hat{\xi} \neq 0$, take $\varepsilon_{k+1} = \frac{1}{2}\varepsilon_k$, $\underline{x}^{(k+1)} = \underline{x}^{(k)}$

$I_{k+1} = I_k$, this iteration is terminated, the next iteration starts at Step 2.

Case 2: If $\hat{\xi} = 0$, solve the following subproblem:

$$
(4) \qquad
\begin{aligned}
&\max \xi \\
&\underline{r}^T \underline{v} \geq \xi \\
&\underline{s}^T \underline{v} \geq \nu\, \xi \qquad \text{if} \quad G(\underline{x}^{(k)}) = p \\
&v_i \geq 0 \qquad \text{if} \quad z_i^{(k)} = 0, \quad k=1,\ldots,n-m. \\
&||\underline{v}|| \leq 1
\end{aligned}
$$

Let the optimal solution be denoted by $(\underline{v}_o, \xi_o)$. Here again we distinguish two cases:

if $\xi_o = 0$ the algorithm terminates, $\underline{x}^{(k)}$ is an optimal solution of problem (1).

if $\xi_o > 0$, the iteration terminates, $\varepsilon_{k+1} = \frac{1}{2}\varepsilon_k$, $\underline{x}^{(k+1)} = \underline{x}^{(k)}$, $I_{k+1} = I_k$ the next iteration starts at Step 2.

Step 4: Let $\underline{v}^{(k)} = \hat{\underline{v}}$, $\underline{u}^{(k)} = -B^{-1} C\, \underline{v}^{(k)}$ and the direction:

$\underline{w}^{(k)} = (\underline{u}^{(k)}, \underline{v}^{(k)})$

Step 5: Compute a starting step length:

$$\alpha_k = \begin{cases} \min\left\{ -\dfrac{x_i^{(k)}}{w_i^{(k)}} \right\} & \text{if there Texists i, } 1 \leq i \leq n, \text{ for which } w_i^{(k)} < 0 \\ \quad w_i^{(k)} < 0 & \text{holds.} \\ R & \text{otherwise} \end{cases}$$

Step 6: Determine the step length λ_k of the iteration as follows:

$$\lambda_k = \frac{1}{2^{\ell_o}} \alpha_k \qquad \text{where} \quad \ell = \ell_o \quad \text{is the first integer, for which}$$

$$(5) \qquad f\left(\underline{x}^{(k)} + \frac{\alpha_k}{2^{\ell}} \underline{w}^{(k)}\right) \geq f(\underline{x}^{(k)}) + \frac{\alpha_k}{2^{\ell+1}} \hat{\xi}$$

$$G\left(\underline{x}^{(k)} + \frac{\alpha_k}{2^{\ell}} \underline{w}^{(k)}\right) \geq p$$

$\ell \geq 0$ if at Step 5 the α_k was computed according to the first row.

holds, i.e. either at Step 5 α_k is computed according to the first row and $\ell = \ell_o = 0$ holds, or system (5) is not fulfilled for $\ell = \ell_o - 1$.

Step 7: $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \underline{w}^{(k)}$

Step 8: If at $\underline{x}^{(k+1)}$ for all $i \in I_k$ $x_i^{(k+1)} > \varepsilon_k$ holds, then this iteration terminates, $\varepsilon_{k+1} = \varepsilon_k$, $I_{k+1} = I_k$, the next iteration starts at Step 1.

Otherwise starting from the present base a new base is determined, such that for the corresponding index set I_{k+1} for every $i \in I_{k+1}$ $x_i^{(k+1)} > \varepsilon_{k+1}$

holds, where $\varepsilon_{k+1} = \frac{1}{2^q} \varepsilon_k$ and either $q=0$ or q is the first positive integer for which a base having the above property can be found. Assumption e.) implies that this is always possible having ε_{k+1} small enough, for a practical procedure the interested reader may consult [4].

For I_{k+1} the matrix B^{-1} is updated accordingly, the iteration terminates and a new iteration starts at Step 1.

In the next part of this section we analyze the algorithm defined above.

In the case of finite termination the current point is an optimal solution of our problem. In fact:

Lemma 2: If the algorithm terminates at the k-th iteration at Step 3 with $\xi_o = 0$, then $\underline{x}^{(k)}$ is an optimal solution of problem (1).

Proof: $\xi_o = 0$ implies that the following problem

$$
\begin{aligned}
&\max \xi \\
&\nabla_{\underline{x}}^T f(\underline{x}^{(k)})\, \underline{w} \geq \xi \\
&\nabla_{\underline{x}}^T G(\underline{x}^{(k)})\, \underline{w} \geq \nu\xi \qquad \text{if} \quad G(\underline{x}^{(k)}) = p \\
&A\,\underline{w} = \underline{0} \\
&v_i \geq 0 \qquad \text{if} \quad z_i^{(k)} = 0, \\
&\qquad\qquad i = 1,\ldots,n-m
\end{aligned} \tag{6}
$$

also has a finite optimum, and at the optimal solution the value of its objective function is 0. Now the same argument works as in [9]. Using the Farkas-Theorem and the fact that assumption c.) guarantees a strictly positive multiplier for the objective function row, from (6) the Kuhn-Tucker optimality conditions can be derived. According to Lemma 1 these are sufficient conditions of the optimality.

In the case of an infinite sequence of iterations the sequence of objective function values increases, furthermore the algorithm continually improves the value of the objective function:

Lemma 3: Either the algorithm terminates at an optimal solution of problem (1), or for an infinite subsequence of the sequence of iterations it reaches Step 4, i.e. at this iterations the value of the objective function is strictly increased.

Proof. Let us observe that the iteration can be terminated at Step 3 only a finite times in sequence. From this fact our assertion follows easily.

Before proceeding let us mentione that there exists a constant $Q > 0$, such that for every iteration $||\underline{w}^{(k)}|| \leq Q$ holds. In fact, because there exist only a finite number of possible bases there exists a $Q > 0$ such that:

$$||\underline{w}^{(k)}|| = ||(-B^{-1} C \underline{v}^{(k)}, \underline{v}^{(k)})|| \leq Q \, ||\underline{v}^{(k)}|| = Q$$

There is one step in the algorithm which requires some discussion, we have namely to prove that at Step 6 it is always possible to select λ_k as required there. This is a trivial consequence of the following statement:

Lemma 4: If $\lambda \leq \frac{1}{2} \frac{\hat{\xi}}{Q^2 L}$, then $f(\underline{x}^{(k)} + \lambda \underline{w}^{(k)}) \geq f(\underline{x}^{(k)}) + \frac{1}{2} \lambda \hat{\xi}$

Proof: Applying the Taylor-Theorem and using the Lipschitz-continuity of $\nabla f(\underline{x})$ we get:

$$f(\underline{x}^{(k)} + \lambda \underline{w}^{(k)}) = f(\underline{x}^{(k)}) + \lambda \nabla_{\underline{x}}^T f(\underline{x}^{(k)}) \underline{w}^{(k)} + \lambda [\nabla_{\underline{x}}^T f(\underline{x}^{(k)} + \Theta \lambda \underline{w}^{(k)}) - \nabla_{\underline{x}}^T f(\underline{x}^{(k)})] \underline{w}^{(k)} \geq$$

$$\geq f(\underline{x}^{(k)} + \lambda \hat{\xi} - \lambda^2 Q^2 L \geq f(\underline{x}^{(k)}) + \frac{1}{2} \lambda \hat{\xi}, \quad \text{if} \quad \lambda \leq \frac{1}{2} \frac{\hat{\xi}}{Q^2 L} \quad \text{where} \quad 0 < \Theta < 1.$$

In the remainder of this section we give a convergence proof for the algorithm. The crucial point of convergence proofs for algorithms which work with

feasible directions and use the same anti zig-zaggig device is to prove that $\varepsilon_k \to 0$ as $k \to \infty$. [11].
This holds also in our case:

Theorem 2: If the algorithm generates an infinite sequence of points $\underline{x}^{(k)}$ then $\varepsilon_k \to 0$ as $k \to \infty$

Proof: Using the same argument as in Lemma 4 if $G(\underline{x}^{(k)}) \leq p + \varepsilon_k$ and a more simple reasoning also based on Taylor's theorem if $G(\underline{x}^{(k)}) > p+\varepsilon_k$ it can be derived that $G(\underline{x}^{(k)}+\lambda\underline{w}^{(k)}) \geq p$ provided that

$\lambda \leq \min\left\{\frac{\nu\hat{\xi}}{Q^2L}, \frac{\varepsilon_k}{K \cdot Q}\right\}$, where K is an upper bound for $||\nabla G(\underline{x})||$.

It is now clear that for $\lambda \leq \min\left\{\frac{1}{2}\frac{\hat{\xi}}{Q^2L}, \frac{\nu\hat{\xi}}{Q^2L}, \frac{\varepsilon_k}{K \cdot Q}, \frac{\varepsilon_k}{Q}\right\}$ the point $\underline{x}^{(k)} + \lambda\underline{w}^{(k)}$ is a feasible solution of problem (1), and in addition

(7) $$f(\underline{x}^{(k)} + \lambda\underline{w}^{(k)}) \geq f(\underline{x}^{(k)}) + \frac{1}{2}\lambda\hat{\xi}$$ holds.

According to Lemma 3 there exists a subsequence for which $\hat{\xi}_k > \varepsilon_k$ is fulfilled, where $\hat{\xi}_k$ denotes $\hat{\xi}$ at the k-th iteration. Let us denote the set of indices in such a subsequence by J. For $k \in J$ the point $\underline{x}^{(k)} + \lambda\underline{w}^{(k)}$ is a feasible solution of problem (1) and (7) holds, if $\lambda \leq \varepsilon_k\varepsilon_o$, where

$$\varepsilon_o = \min\left\{\frac{1}{2Q^2L}, \frac{\nu}{Q^2L}, \frac{1}{KQ}, \frac{1}{Q}\right\}$$

At Step 6 of the algorithm ℓ_o was the first integer, for which (5) holds, so necessarily $\lambda_k > \frac{1}{2}\varepsilon_k\varepsilon_o$, for $k \in J$. Substituting into (5) we get:

$$f(\underline{x}^{(k+1)}) \geq f(\underline{x}^{(k)}) + \frac{1}{4}\varepsilon_o\varepsilon_k^2 \quad , \quad k \in J$$

Since $f(\underline{x}^{(k)})$ is a monotone increasing sequence bounded from above, it follows that $\varepsilon_k \to 0$, $(k\to\infty)$, $k \in J$. However ε_k is a monotone decreasing sequence so $\varepsilon_k \to 0$, $(k\to\infty)$ holds.

Based on this theorem, the convergence of the algorithm can be proved.

Theorem 3: Either the algorithm terminates in a finite number of iterations at an optimal solution of *(1)*, or every accumulation point of the sequence $\underline{x}^{(k)}$, $k = 1,2,\ldots$ is an optimal solution of problem *(1)*.

Proof: Because of the fact that $f(\underline{x})$ is a concave function and $f(\underline{x}^{(k)})$, $k = 1,2,\ldots$ a monotone sequence, it is sufficient to prove that in the case of an infinite sequence there exists a convergent subsequence of the sequence $\underline{x}^{(k)}$, $k = 1,2,\ldots$ with a limit point being an optimal solution of *(1)*. To select such a subsequence first of all we mention that assumption e.) implies that ε_k is reduced at Step 8 only a finite times. Using this remark from Theorem 2 it follows that there exists a subsequence of the sequence of iterations, for which $\hat{\xi}_k \leq \varepsilon_k$, $k \in J$ holds, where J denotes the set of indices in this subsequence. According to assumption d.) we may assume that the sequence $\underline{x}^{(k)}$, $k \in J$ converges to some feasible point $\underline{x}^*$, i.e. $\underline{x}^{(k)} \to \underline{x}^*$, $(k\to\infty)$, $k \in J$. We consider the following problem:

$$
\begin{aligned}
&\max \xi \\
&\nabla_{\underline{x}}^T f(\underline{x}^*)\, \underline{w} \geq \xi \\
(8)\qquad &\nabla_{\underline{x}}^T G(\underline{x}^*)\, \underline{w} \geq \nu\xi \qquad \text{if} \quad G(\underline{x}^*) = p \\
&A\underline{w} = \underline{0} \\
&w_i \geq 0 \qquad \text{if} \quad x_i^* = 0\,, \quad i = 1,\ldots,n. \\
&||\underline{w}|| \leq 1
\end{aligned}
$$

Let us denote the optimal solution of this problem by $(\underline{w}^*, \xi^*)$.

If here $\xi^*=0$ holds, then using the same technique as in the proof of Lemma 2 it follows easily that $\underline{x}^*$ is an optimal solution of problem *(1)*.

Let us assume that $\xi^* > 0$, we shall get a contradiction. If $G(\underline{x}^*) > p$, then for k large enough $G(\underline{x}^{(k)}) > p + \varepsilon_k$, $k \in J$. Similarly if for some i, $1 \leq i \leq n$, $x_i^* > 0$ holds then for k large enough $x_i^{(k)} > \varepsilon_k$, $k \in J$ holds. On the other hand because of the continuity of $\nabla f(\underline{x})$ and $\nabla G(\underline{x})$ for k sufficiently large

$$\nabla_{\underline{x}}^T f(\underline{x}^{(k)})\ \underline{w}^* \geq \frac{1}{2}\ \xi^*$$
$$\nabla_{\underline{x}}^T G(\underline{x}^{(k)})\ \underline{w}^* \geq \frac{1}{2}\ \nu\xi^* \qquad \text{if} \qquad G(\underline{x}^*) = p$$

holds. Let us further consider all bases of A, and denote by t the maximum of the norms of the corresponding non-basic parts of $\underline{w}^*$. From our considerations it follows that the non-basic parts of $(\frac{1}{t}\ \underline{w}^*, \frac{1}{2t}\ \xi^*)$ form a feasible solution to the direction-finding subproblems (3), for $k \geq k_o$, $k \in J$ where k_o is sufficiently large. This implies that $\hat{\xi}_k \geq \frac{1}{2t}\ \xi^*$ for $k \geq k_o$, $k \in J$ which contradicts the fact that because $\hat{\xi}_k \leq \varepsilon_k$ for $k \in J$, the sequence $\hat{\xi}_k$ tends to 0 for $k \to \infty$, $k \in J$. This completes the proof.

Some final comments:

If we choose the norm $||\underline{v}|| = \max_{1 \leq i \leq n-m} |v_i|$, then the subproblems (3), (4) (8) are linear programming problems with only two rows, they can easily be solved. The method can be extended for more than one nonlinear constraint, if we have ℓ nonlinear constraints, the above mentioned LP sub-problems contain $\ell+1$ rows. This suggests that the method might work well with only a few nolinear constraints.

The method reduces to the reduced gradient method if there are only linear constraints, and it gives Zoutendijk's P1 method if except of bounds for the variables there are nonlinear constraints.

The first phase of the method, that is the determination of a starting feasible point $\underline{x}^{(1)}$ and the corresponding base is carried out by maximizing $G(\underline{x})$ subject to the linear constraints, like in [10].

The idea of modifying the reduced gradient method in order to work with feasible points was first proposed by Kleinmichel and Sadowski [3], [12], but they

use different approach and get a different algorithm. Kleinmichel also gave a nice and general framework for proving the convergence of methods which operate with feasible directions [3].

5. Computational experiences with the method.

The numerical difficulties, which arise in the implementation of the reduced gradient method for problem (1) are already explained in Section 2. Here we only mention that the method described in the previous section proved to be superior to the specialized version of the reduced gradient method for all of our stochastic programming test problems.

First the algorithm was tested on a series of small-size problems, with satisfactory results. For instance the computational results are in accordance with a theorem of Slepian, which states that the two dimensional normal probability distribution function $\Phi(x_1, x_2; r)$ is a monotone increasing function of the correlations coefficient r, at any fixed (x_1, x_2) point.

Having thoroughly tested the algorithm we also solved a version of a model of the electric energy sector of the Hungarian Economy, a real-life problem. The detailed description of this model can be found in the paper of A.Prékopa, S. Ganczer, I. Deák, K. Patyi [10], together with a thorough analysis of Zoutendijk's P2 method for this case, and some very interesting computational results. The model is a stochastic programming problem as described in Section 2, with 49 linear constraints and 72 variables after introducing slack variables for some inequality constraints. The stochastic constraint contained 2 rows. We solved the problem with $p = 0.90$ and $p = 0.95$, and have got essentially the same optimum as reported in [10].

The code was developed on the framework given in the perivous section, in FORTRAN language, for a CDC 3300 computer. Some computational results of the above mentioned two runs are summarized in the following tableau:

	Phase I.	Phase II.	
		p=0.90	p=0.95
Starting value of the objective	0.49993	3875.00	3875.00
Optimal value of the objective	0.99643	4371.68	4370.73
Value of $G(\underline{x})$ at the optimum	0.99643	0.89999	0.95001
Number of iterations	19	18	16
Number of the calculations of $G(\underline{x})$	37	68	61
Computing time in seconds	51.24	45.71	40.82

Although we are aware of its insufficiency, just to give some comparison we mention that CDC's LP package REX needed 51 seconds to solve in two phases the deterministic version of the problem on the CDC 3300.

Acknowledgement. I wish to express my thanks to Professor A. Prékopa for his advices and for suggesting this research.

REFERENCES

[1] J. Abadie, J. Carpentier, "Généralisation de la méthode du gradient réduit de Wolfe au cas des contraintes non-linéaires" in.Proc. IFORS Conf., eds. D.B. Hertz and J. Melese, Wiley, New York, 1966, 1041-1053.

[2] I. Deák, "Monte-Carlo evaluation of the multidimensional normal distribution function by the ellipsoid method", Problems of Control and Information Theory, to appear.

[3] H. Kleinmichel, "On methods of feasible directions", Lecture at III. Conf. on Math. Progr., Mátrafüred, 1975.

[4] H. Kleinmichel, H. Sadowski, "Der verallgemeinerte RG-Algorithmus bei linearen Restriktionen, die Behandlung des Entartungsfalls und die Konvergenz des Verfahrens", Beiträge zur Num. Math., 3/1975/,37-55.

[5] J. Mayer, "Computational experiences with the reduced gradient method", in: Progress in Op. Res., ed. A. Prékopa, North- Holland, 1974, 613-624.

[6] A. Prékopa, "On probabilistic constrained programming", in: Proc. of the Princeton Symposium on Mathematical Programming, Princeton Univ. Press, Princeton, N.Y., 1970, 113-138.

[7] A. Prékopa, "Logarithmic concave measures with application to stochastic programming", Acta Sci. Math., Szeged, 32/1971/, 301-316.

[8] A. Prékopa, "Contributions to the theory of stochastic programming", Mathematical Programming, 4/1973/, 202-221.

[9] A. Prékopa, "Eine Erweiterung der sogenannten Methode der Zulässigen Richtungen der nichtlinearen Optimierung auf den Fall quasikonkaver Restriktionen", Math. Operationsforsch. u. Statist., 5/1974/, Heft 4/5, 281-293.

[10] A. Prékopa, S. Ganczer, I. Deák, K. Patyi, "The STABIL stochastic programming model and its experimental application to the electric energy sector of the Hungarian Economy", Proc. of the Intern. Conf. on Stochastic Programming, Oxford, 15-17 July, 1974. Academic Press, in print.

[11] B.N. Pshenichnyj, JU.M. Danilin, "Computational methods in optimization", Izd. Nauka, 1975. in Russian.

[12] H. Sadowski,"Untersuchungen zu den Verfahren der reduzierten Gradienten in der nichtlinearen Optimierung", Ph.D. Thesis, University of Technology, Dresden, GDR, 1973.

[13] P. Wolfe, "Methods of nonlinear programming", in: Recent advances in mathematical programming, eds. R.L. Graves and P.Wolfe, McGraw-Hill, New York, 1963, 76-77.

[14] G. Zoutendijk, "Methods of feasible directions", Elsevier Publ. Comp., Amsterdam and New York, 1960.

RELIABILITY TYPE INVENTORY MODELS BASED ON STOCHASTIC PROGRAMMING

András Prékopa

TECHNOLOGICAL UNIVERSITY and COMPUTER AND AUTOMATION INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST, HUNGARY

Péter Kelle

COMPUTER AND AUTOMATION INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST, HUNGARY

ABSTRACT

The models discussed in the present paper are generalizations of the models introduced previously by A. Prékopa [6] and M. Ziermann [13]. In the mentioned papers the initial stock level of one basic material is determined where the delivery and demand processes allow certain homogenity (in time) assumptions if they are random. Here we are delaing with more than one basic materials and drop the time homogenity assumption. Only the delivery processes will be assumed to be random. They will be supposed to be stochastically independent. Out of the models discussed in this paper the first one was already introduced in [9]. All these models are stochastic programming models where algorithms serve for the determination of the initial stock levels instead of simple formulas. We have to solve nonlinear programming problems where one of the constraints is probabilistic. The function and gradient values of the corresponding constraining function are determined by simulation. Numerical example is given.

1. Introduction. The models didcussed in the present paper are generalizations of the models intrudoced previously by A. Prékopa [6] and M. Ziermann [13]. In the mentioned papers the initial stock level

of one basic material is determined where the delivery and demand process allow certain homogenity (in time) assumptions if they are random. Here we are dealing with more than one basic materials and drop the time homogenity assumption. Only the delivery processes will assumed to be random. They will be supposed to be stochastically independent. Out of the models discussed in this paper the first one was already introduced in [9]. All these models are stochastic programming models where algorithms serve for the determination of the initial stock levels instead of simple formulas. We have to solve nonlinear programming problems where one of the constraints is probabilistic. The function and gradient values of the corresponding constraining function are determined by simulation.

The numerical evaluation of the models discussed in the present paper is more sophisticated than those of the earlier models of Prékopa and Ziermann. However, if the delivery process is inhomogeneous then with the present methodology we can get closer the reality and can handle many delivery processes simultaneously.

The most general model introduced in [6] is the following. Let M denote the initial stock level, $(0,T)$ the investigated time interval, α_t the amount of the basic material delivered up to time t and β_t the cumulative demand which occured up to t, where $0 \leq t \leq T$. The initial stock level is to be determined in such a way that this be smallest M satisfying

$$(1.1) \qquad P\left(\inf_{0 \leq t \leq T} (M + \alpha_t - \beta_t) > 0\right) \geq 1 - \varepsilon$$

where ε is a previously prescribed, in practice low value, e.g. $\varepsilon=0,05$. Under the assumptions introduced in [6] in connection with the random processes α_t, β_t Relation (1,1) holds with equality in case of the optimal initial stock. Thus an equation serves for the determination of M. This is called the Reliability Equation.

For the easier understanding of the generalizations we present in this paper, we need to repeat the modelling of the random processes α_t, β_t

introduced in [6]. Since the model for β_t is the same as that of α_t only its parameters are different, it will be enough to deal with α_t only.

Let λ be a real number satisfying $0 \leq \lambda \leq 1$ and $t_1,\dots,t_n$ further $\tau_1,\dots,\tau_{n-1}$ be independent samples taken from the population uniformly distributed in $(0,1)$. Let $\tau_1^* \leq \tau_2^* \leq \dots \leq \tau_{n-1}^*$ be the ordered sample corresponding to τ_i, $i = 1,\dots,n-1$ and put $\tau_o^* = 0$, $\tau_n^* = 1$. Now we define α_t in the following manner:

$$\alpha_t = c\lambda\nu/n + c(1-\lambda)\tau_\nu^*, \quad 0 \leq t \leq T, \tag{1.2}$$

where ν is the number of those t_i which are smaller than t, c is a positive constant, cT equals the total demand occuring in the time interval $(0,T)$ and this is supposed to be equal the total amount of basic material delivered in the same time interval. If $\lambda = 1$ then α_t is the empirical probability distribution function belonging to the sample $t_1,\dots,t_n$. In connection with β_t we use m instead of n and μ instead of λ.

In [6] it is proved that the following limit relations hold:

(1.3)

$$\lim_{\substack{m\to\infty \\ n\to\infty}} P\left(\left(\frac{mn}{m+n+m(1-\lambda)^2+n(1-\mu)^2}\right)^{\frac{1}{2}} \sup_{0\leq t\leq 1} (\alpha_t - \beta_t) < y\right) =$$

$$= \lim_{\substack{m\to\infty \\ n\to\infty}} P\left(\left(\frac{mn}{m+n+m(1-\lambda)^2+n(1-\mu)^2}\right)^{\frac{1}{2}} \sup_{0\leq t\leq 1} (\beta_t - \alpha_t) < y\right) =$$

$$= \begin{cases} 1 - \exp(-2y^2), & \text{if } y > 0 \\ 0, & \text{if } y \leq 0. \end{cases}$$

Here we have fixed $T=1$ for the sake of simplicity. This choice does not restrict the generality.

If we assume the left hand sides of *(1.3)* approximately equal to the right hand side value then for a given ε the following $M=M_{\lambda,\mu}$ value turns out to be the approximate solution of the Reliability Equation:

$$(1.4) \qquad M_{\lambda,\mu} = c\left[\frac{1}{2}\left(\frac{1+(1-\lambda)^2}{n} + \frac{1+(1-\mu)^2}{m}\right)\log\frac{1}{\varepsilon}\right]^{\frac{1}{2}}$$

If α_t is a deterministic process and $\alpha_t = ct$ $(0 \leq t \leq 1)$, then the corresponding M value can be ontained from *(1.4)* if we take the limit $n \to \infty$. We proceed similarly if β_t is deterministic. We remark that the minimal amount δ delivered at one delivery time and λ are in the following relation: $\lambda = n\delta/c$. Similar relation holds for the parameters of the process β_t.

2. Generalization of the delivery and demand processes.

In this section we repeat the generalization fo the delivery process as it is given in [9].

In Section 1. we mentioned the following assumptions in connection with the delivery process:

a) the number of delivery times is fixed, this was denoted by n;

b) the n delivery time points are so distributed in the interval (0,1) as the elements of a sample of size n taken from a population uniformly distributed in the same interval;

c) the total delivered amount is constant and is equal to c which is at the same time equal to the total demand;

d) the random vector the components of which are the random delivered amounts is stochastically independent of the random vector of the delivery time points;

e) denoting by δ the smallest amount to be delivered if one delivery occurs, the model for the distribution of the remaining amount among the n delivery times is the following: divide the interval $(0, c - n\delta)$ into n parts by choosing $n-1$ independent and uniformly distributed random points and assign the quantitites equal to the lengths of the subintervals to the n delivery times.

In what follows we maintain the assumptions a), c), d), and modify the assumptions b),e). For the modelling of the delivery process we choose L uniformly distributed independent random points in the interval $(0, c-n\delta)$, where $L > n-1$. Let $y_1^*, \ldots, y_L^*$ denote the ordered sample formed from the L random points. Out of this ordered sample we select those which have subscripts $k_1 < k_2 < \ldots < k_{n-1}$ and add to the fixed delivery amounts the following

$$(2.1) \qquad \eta_1 = y_{k_1}^* , \quad \eta_2 = y_{k_2}^* - y_{k_1}^* , \ldots, \quad \eta_n = c - n\delta - y_{k_{n-1}}^* .$$

Thus the amounts delivered at the delivery times will be

$$(2.2) \qquad \delta + \eta_1 , \quad \delta + \eta_2, \ \ldots, \ \delta + \eta_n .$$

Similar model is used for the delivery time points. To the fixed amount to be delivered at one occasion there corresponds a fixed time γ as the minimal distance between two consecutive delivery time points $(0 \leq \gamma \leq 1/n)$. The delivery time points are selected from an ordered sample $x_1^* \leq x_2^* \leq \ldots \leq x_N^*$ of a sample of size N taken from a population uniformly distributed in $(0, 1 - n\gamma)$, so that we select those elements which have subscripts $j_1 < j_2 < \ldots < j_n$, for the random variables

$$(2.3) \qquad \xi_1 = x_{j_1}^* , \quad \xi_2 = x_{j_2}^* - x_{j_1}^* , \ldots, \quad \xi_n = x_{j_n}^* - x_{j_{n-1}}^* .$$

and finally take the partial sums of the random variables

$$(2.4) \qquad \gamma + \xi_1 \ , \ \gamma + \xi_2 \ , \ldots , \ \gamma + \xi_n .$$

This partial sums represent the n delivery time points.

Let $s(z_1 , \ldots , z_{n-1})$ denote the joint probability density function of the random variables $\eta_1 , \ldots , \eta_{n-1}$. It is easy to see that this function has the following form

$$s(z_1 , \ldots , z_{n-1}) =$$

$$= \left(\frac{1}{c-n\delta} \right)^n \frac{\Gamma (L+1)}{\Gamma(k_1)\Gamma(k_2-k_1)\ldots\Gamma(k_{n-1}-k_{n-2})\Gamma(L+1-k_{n-1})} \cdot$$

$$(2.5) \qquad \cdot \left(\frac{z_1}{c-n\delta} \right)^{k_1-1} \left(\frac{z_2}{c-n\delta} \right)^{k_2-k_1-1} \ldots .$$

$$\ldots . \left(\frac{z_{n-1}}{c-n\delta} \right)^{k_{n-1}-k_{n-2}-1} \left(1 - \frac{z_1 + \ldots + z_{n-1}}{c-n\delta} \right)^{L-k_{n-1}}$$

if $z_i > 0$, $i=1,\ldots,n-1$; $z_1+\ldots+z_{n-1} < c-n\delta$ and $s(z_1,\ldots,z_{n-1}) = 0$ otherwise.

Similar formula gives the joint probability density function of the random variables $\xi_1 , \ldots , \xi_n$ which we denote by $r(z_1 , \ldots , z_n)$.

$$r(z_1, \ldots, z_n) =$$

$$= \left(\frac{1}{1-n\gamma} \right)^{n+1} \frac{\Gamma(N+1)}{\Gamma(j_1)\Gamma(j_2-j_1) \ldots \Gamma(j_n-j_{n-1})\Gamma(N+1-j_n)} \cdot$$

(2.6)

$$\cdot \left(\frac{z_1}{1-n\gamma} \right)^{j_1-1} \left(\frac{z_2}{1-n\gamma} \right)^{j_2-j_1-1} \ldots$$

$$\ldots \left(\frac{z_n}{1-n\gamma} \right)^{j_n-j_{n-1}-1} \left(1 - \frac{z_1 + \ldots + z_n}{1-n\gamma} \right)^{N-j_n}$$

if $z_i > 0$, $i = 1, \ldots, n$; $z_1 + \ldots + z_n < 1 - n\gamma$ and $r(z_1, \ldots, z_n) = 0$ otherwise.

Thus the random vectors $(\eta_1, \ldots, \eta_{n-1})$ and $(\xi_1, \ldots, \xi_n)$ have Dirichlet distributions. For properties of this multivariate distribution the reader is referred to [12].

3. The Inventory Models.

Model I. [9]. The model for the delivery process is that one discussed in Section 2. The demand is assumed to have constant intensity i.e. the demand occuring in the interval $(0,t)$ is equal to ct where c is a constant. M denotes the initial stock level. The demand will be met continuously in the whole interval $(0,1)$ if and only if the following relations hold

$$
\begin{aligned}
&M && \geq \gamma + \xi_1 \\
&M + \delta + \eta_1 && \geq 2\gamma + \xi_1 + \xi_2 \\
&M + 2\delta + \eta_1 + \eta_2 && \geq 3\gamma + \xi_1 + \xi_2 + \xi_3 \\
& && \vdots \\
&M + (n-1)\delta + \eta_1 + \ldots + \eta_{n-1} && \geq n\gamma + \xi_1 + \xi_2 + \ldots + \xi_n
\end{aligned}
\tag{3.1}
$$

Let us introduce the notations:

$$
\begin{aligned}
\zeta_1 &= \xi_1 \\
\zeta_2 &= \xi_1 + \xi_2 - \eta_1 \\
&\vdots \\
\zeta_n &= \xi_1 + \ldots + \xi_n - \eta_1 - \ldots - \eta_{n-1}
\end{aligned}
\tag{3.2}
$$

The random vectors $\underline{\eta} = (\eta_1, \ldots, \eta_{n-1})$ and $\underline{\xi} = (\xi_1, \ldots, \xi_n)$ are independent and their probability density functions are logconcave functions in R^{n-1} resp. R^n. It follows that the altogether $2n-1$ components have a logconcave joint density in R^{2n-1}.

The notation of a logconcave probability measure is introduced in [7]. A **probability measure** P defined on the measurable subsets of R^m is said to be logconcave if for every pair A,B of convex subsets, of R^m and every $0 < \lambda < 1$ the following inequality holds:

$$
P(\lambda A + (1-\lambda)B) \geq [P(A)]^{\lambda} [P(B)]^{1-\lambda} \tag{3.3}
$$

The main theorem of [7] says that if a probability measure is generated by a logconcave probability density, then it is a logconcave measure. On the other hand any linear transform of a random vector having logconcave distribution has again logconcave distribution [9, Theorem 3]. Thus the random vector $\underline{\zeta} = (\zeta_1, \ldots, \zeta_n)$ has logconcave probability distribution.

We can write the Reliability Equation in our case, by taking into account only one basic material, as follows:

$$h(M) = P\,(\zeta_i \leqq M + (i-1)\delta - i\gamma, \quad i=1,\ldots,n) = p\,, \tag{3.4}$$

where $0 < p < 1$ and $p \approx 1$ in the practice.

The function $h(M)$ is logconcave on the halfline $[0,\infty]$ bacause the joint probability distribution function of a random vector having logconcave probability distribution is a logconcave point function [7]. Model I for more than one basic material consists of the following stochastic programming problem:

$$d^{(1)}\, M^{(1)} + \ldots + d^{(\ell)} M^{(\ell)}$$

is to be minimized supposing that

$$\begin{aligned} &h\,(\underline{M}) = h_1(M^{(1)}) \ \ldots \ h_\ell(M^{(\ell)}) \geqq p \\ &\underline{M} \geqq \underline{0} \quad , \quad \underline{M} \in D, \end{aligned} \tag{3.5}$$

where $\underline{M} = (M^{(1)}, \ldots, M^{(\ell)})$ and D is a subset of R^ℓ determined by some constraints such as we prescribe that the components of $\underline{M}$ be smaller than ot equal to certain upper bounds or that the initial stock amounts do not take more room than a certain upper limit and do not require more financial investment than a further upper limit etc. The numbers $d^{(1)}, \ldots, d^{(\ell)}$ are nonnegative and they are some

valuations of units of goods to be determined on the basis of local knowledge. Sometimes nonlinear objective function may turn out.

In the above discussion we assumed the demand function to be linear. Of course we can drop this assumption and use that model for the demand processes what was introduced in Section 2.

Stochastic programming models with independent joint constraints were considered first by Miller and Wagner [4].

Model II. This model differs from the previous one in that further constraints containing conditional expectation appear. With this we prescribe not only the rarity of the occurance of unsatisfied demand but also prescribe upper bound for the average magnitude of the unsatisfied demand. Thus upper bound may depend on the basic material. We assume that unsatisfied demand will not be lost. Thus the model works with backorders.

If one of the inequalities (3.1) is violated, then it means that there was a lack just before the considered delivery time. As large is the violation as great is the unsatisfied demand. Assuming deterministic demand process with constant intensity, this means that the length of the time interval in which unsatisfied demand existed, is proportional to the magnitude of the violation. Here did we take into account that no demand is lost. Our model consists of the problem formulated below. The superscripts refer to the various basic materials:

$d^{(1)}M^{(1)} + \ldots + d^{(\ell)} M^{(\ell)}$ is to be minimized supposing that

$$h(\underline{M}) \geq p,$$

$$(3.6) \qquad E(\zeta_i^{(j)} - M^{(j)} - (i-1)\delta^{(j)} + i\gamma^{(j)} \mid \zeta_i^{(j)} - M^{(j)} + (i-1)\delta^{(j)} + i\gamma^{(j)} > 0) \leq g_i^{(j)}$$

$$i = 1, \ldots, n \quad , \quad j = 1, \ldots, \ell,$$

$$\underline{M} \geq \underline{0} \quad , \quad \underline{M} \in D \, ,$$

where the $g_i^{(j)}$ are constants and E is the symbol of expectation. The conditional expectation type constaints may even substitute the

the probabilistic constraint.

For every i,j the random variable $\zeta_i^{(j)}$ has a logconcave probability density. It follows from this (see [8]) that the constraining functions in the conditional expectation constraints are monotonically decreasing functions of the variables $M^{(j)}$ and every such constraint is simply equivalent to a lower bound for the variable $M^{(j)}$ appearing in the constraint. We return to this question at the end of the section.

Model III. Again we assume that no demand will be lost. The difference between this model and (3.6) consists in a penalty term what we introduce now. Let us introduce the random variables

$$(3.7)\quad k_i^{(j)} = \begin{cases} q_i^{(j)} \cdot (\zeta_i^{(j)} - M^{(j)} - (k-1)\delta^{(j)} + i\gamma^{(j)}) \\ \quad \text{if} \quad \zeta_i^{(j)} - M^{(j)} - (i-1)\delta^{(j)} + i\gamma^{(j)} > 0 \\ 0 \qquad \text{otherwise} \end{cases}$$

$i=1,\ldots,n$; $j=1,\ldots,\ell$ where $q_i^{(j)} \geq 0$ for every i and j. It is easy to show that $E(k_i^{(j)})$ is a convex function of the variable $M^{(j)}$. To this it is enough to know that $\zeta_i^{(j)}$ has a continuous probability distribution. Since Model III. has only a new objective function as compared to the model given by (3.6), it will be enough to formulate the new objective function. This is the following:

$$(3.8)\qquad \Sigma_{j=1}^{\ell} \; d^{(j)} \; M^{(j)} + \Sigma_{j=1}^{\ell} \; \Sigma_{i=1}^{n} \; E(K_i^{(j)})$$

The construction of the above three models are in correspondence with the three general model constructions given in [8].

Model III. contains Model I. and Model II. as special cases. We obtain Model II. by putting $g_i^{(j)} = 0$, and Model I. by putting $g_i^{(j)} = \infty$, $i = 1,\ldots, n$; $j = 1, \ldots, \ell$.

Now for a while we return to the conditional expectation contained in Problem (3.6) and the expectations $E(k_i^{(j)})$. Let $f_i^{(j)}$ resp. $F_i^{(j)}$ denote the probability density and the probability distribution functions of the random variable $\zeta_i^{(j)}$. For the sake of simplicity the superscripts will be ommitted in the sequel. It is easy to see that if ζ is a continuously distributed random variable and a is a constant, then the following equality holds

$$
\begin{aligned}
E(\zeta-a \mid \zeta-a>0) &= \int_a^{\infty} [1-F(x)]dx/[1-F(a)] \\
&= \int_a^{\infty} xf(x)\,dx\,/[1-F(a)]-a
\end{aligned}
\tag{3.9}
$$

where F is the probability distribution function of ζ. In view of this we can write

$$
\begin{aligned}
&E(\zeta_i-M-(i-1)\delta+i\gamma \mid \zeta_i-M-(i-1)\delta+i\gamma > 0) = \\
&\frac{1}{1-F_i(M+(i-1)\delta-i\gamma)} \int_{M+(i-1)\delta-i\gamma}^{1-n\gamma} xf_i(x)dx-M-(i-1)\delta+i\gamma .
\end{aligned}
\tag{3.10}
$$

Similarly we obtain

$$
E(k_i) = q_i \int_{M+(i-1)\delta-i\gamma}^{1-n\gamma} [1-F_i(x)]\,dx. \tag{3.11}
$$

A simple argument shows that

$$
f_i(x) = \frac{\Gamma(N+1)\,\Gamma(L+1)}{\Gamma(j_i)\Gamma(N-j_i+1)\Gamma(k_{i-1})\Gamma(L-k_{i-1}+1)(c-n\delta)(1-n\gamma)} \cdot
$$

$$
\int_o^b \left(\frac{x+u}{1-n\gamma}\right)^{j_i-1} \left(1-\frac{x+u}{1-n\gamma}\right)^{N-j_i} \left(\frac{u}{c-n\delta}\right)^{k_{i-1}-1} \left(1-\frac{u}{c-n\delta}\right)^{L-k_{i-1}} \tag{3.12}
$$

(where $b = \min\{1-n\gamma-x,\ c-n\delta\}$),

if $0 < x < 1-n\gamma$ and $f_i(x) = 0$ otherwise , $i = 2,3,\ldots,n$,
further

$$f_1(x) = \frac{\Gamma(N+1)}{\Gamma(j_1)\Gamma(N-j_1+1)} \frac{1}{1-n\delta} \left(\frac{x}{1-n\gamma}\right)^{j_1-1} \left(1 - \frac{x}{1-n\gamma}\right)^{N-j_1} \tag{3.13}$$

if $0 < x < 1-n\gamma$ and $f_1(x) = 0$ otherwise.

As we already remarked, the ith conditional expectation-type constraint in Problem (3.6) can be converted into the simple inequality $M^{(j)} \geqq M_i^{(j)}$ where $M_i^{(j)}$ is that value of $M^{(j)}$ for which the constraint hodls with equality. This value can be determined by numerical intergration of the function f_j.

4. SOLUTION OF THE PROBLEMS. In this section we present a solution method to the problems discussed in the previous section. We restrict ourselves to the problem of Model I, since the solution of the two further problems requires only slight modification.

For the sake of simplicity let us agree that the constraint $\underline{M} \in D$ be specialized so that it consist in the system of inequalities $M^{(j)} \leqq 1$, $j = 1,\ldots, \ell$. These are, on the other hand, no real restrictions, because the equalities

$$h_j(1) = 1 \quad , \quad j = 1,\ldots, \ell \tag{4.1}$$

hold trivially and these imply that the optimal $M^{(1)}, \ldots, M^{(\ell)}$ values are automatically smaller than or equal to 1. The upper bounding of the $M^{(1)}, \ldots, M^{(\ell)}$ values has the only significance that we shall be able to refer to well-known convergence theorem concerning the SUMT method what we are going to use here [2].

We apply the interior point version of the SUMT. Consider the following penalty function.

$$G(r,\underline{M}) = \sum_{j=1}^{\ell} d^{(j)}M^{(j)} - r\{\log[\prod_{j=1}^{\ell} h_j(M^{(j)}) - p] + \sum_{j=1}^{\ell} \log M^{(j)}(1-M^{(j)})\} \tag{4.2}$$

where r is a fixed positive number. It is easy to see that for every fixed $p > 0$ the function $h_1(M^{(1)}) \ldots h_\ell(M^{(\ell)})-p$ is also logconcave which implies that for every fixed $r > 0$ the function $G(r,\underline{M})$ is convex on the set $\{\underline{M} \mid \underline{M} \geq \underline{0}\}$. From this we only need the fact that $G(r,M)$ is convex on the ℓ-dimensional unit cube $\{\underline{M} \mid 0 \leq M^{(j)} \leq 1, j = 1,\ldots, \ell\}$. The SUMT interior point method works so that we take a sequence $r_1 > r_2 > \ldots$ consisting of positive numbers, tending to 0 and minimize $G(r_k,\underline{M})$ with respect to $\underline{M}$ (in principle) for every r_k. If $\underline{M}_k$ is the minimizing vector then $G(r_k,\underline{M}_k)$ tends to the minimum value of the objective function in Model I. Thus $\underline{M}_k$ is an approximate optimal solution to the problem if k is large enough.

The mentioned convergence is ensured if the set of feasible solutions is bounded, the constraining functions as well as the objective function are continuous on the set of feasible solutions, further there exists interior point of this set and all inequalities hold as strict inequalities at every interior point. As regards Model 1 here we have the constraints $0 \leq M^{(j)} \leq 1$, $j=1,\ldots,\ell$ and the probabilistic constraint which restricts further the unit cube. The former ones hold strictly at every interior point of the unit cube thus it will be enough to consider the probabilistic constraint. Let $\underline{M}_1$ be an interior point in the set of feasible solutions. We shall show that $h(\underline{M}_1) > p$ where p is fixed and $0 < p < 1$ the line section connecting the points $\underline{1}$ and $\underline{M}_1$ is entirely feasible because the set of feasible solutions is convex. Let $\underline{M}_o$ be a feasible point on the line connecting $\underline{1}$ and $\underline{M}_1$ lying outside the section so that $\underline{M}_1$ be between $\underline{1}$ and $\underline{M}_o$. Then there exists a $0 < \lambda < 1$ so that

$$\underline{M}_1 = \lambda\underline{1} + (1 - \lambda)\underline{M}_o$$

from which, using the logconcavity of the function $h(\underline{M})$, it follows that

$$(4.3) \qquad h(\underline{M}_1) \geq [h(\underline{1})]^{\lambda}[h(\underline{M}_o)]^{1-\lambda} \geq p^{1-\lambda} > p.$$

Thus we have shown that the SUMT interior point method is convergent in case of Model I.

Many general unconstrained minimization technique can be applied for the function (4.2). Some of them use only function values, some use gradient values too. On order to facilitate the application of methods belonging to the latter category, we present a methodology to compute the gradient values. Since $h(\underline{M})$ is the product of the functions $h_j(M^{(j)})$, $j = 1,\ldots,\ell$ it will be enough to consider the derivatives of the functions $h_j(M^{(j)})$. Let us ommit the j, for the sake of simplicity. The function (3.4) is the joint probability distribution function of the random variables $\zeta_1,\ldots,\zeta_n$ at the point with coordinates $M + (i-1)\delta - i\gamma$, $i=1,\ldots,n$.

We remark that if $F(\underline{z})$ is the probability distribution function corresponding to a continuous probability distribution, then the following relation holds:

$$(4.4) \qquad \frac{\partial F(\underline{z})}{\partial z_i} = F(z_j, \quad j\neq i \mid z_i)\, f_i(z_i), \quad i = 1,\ldots,n,$$

where $f_1,\ldots,f_n$ are the probability density functions of the one-dimensional marginal distributions and $F(.\mid z_i)$ is the n-1-dimensional conditional probability distribution function given that the ith random variables equals z_i.

We assume that $n\delta < c$, $n\gamma < 1$ (if one of the equalities $n\delta = c$, $n\gamma = 1$ holds, our procedure can essentially be simplified). To compute the derivative of the function (3.4), first we take partial derivatives with respect to all $z_1,\ldots,z_n$ of the funciton

$$(4.5) \qquad P(\zeta_i \leqq z_i + (i-1)\delta - i\gamma, \quad i = 1,\ldots,n)$$

and put $z_1 = \ldots = z_n = M$. The sum of these equals the derivative of $h(M)$. The partial derivative of the function (4.5) with respect to z_i can be obtained by using the formula (4.4). Putting $z_1 = \ldots z_n = M$ we obtain

$$P\,(\zeta_j \leqq M+(j-1)\delta-j\gamma, \quad j\neq i \mid \zeta_i = M+(i-1)\delta-i\gamma).f_i(M+(i-1)\delta-i\gamma) =$$

$$= f_i(M+(i-1)\delta-i\gamma)\cdot \int_0^v P(\zeta_j \leqq M+(j-1)\delta-j\gamma,$$

$$j \neq i \mid \xi_1 +\ldots+ \xi_i = M+(i-1)\delta-i\gamma + x,$$

$$\eta_1 +\ldots+ \eta_{i-1} = x)\ \frac{\Gamma(N+1)}{\Gamma(j_i)\Gamma(N+1-j_i)} \quad \frac{1}{1-n\gamma}\left(\frac{M+(i-1)\delta-i\gamma+x}{1-n\gamma}\right)^{j_i-1} \cdot$$

$$\cdot \left(1 - \frac{M+(i-1)\delta-i\gamma+x}{1-n\gamma}\right)^{N-j_i} \quad \frac{\Gamma(L+1)}{\Gamma(k_{i-1})\Gamma(L+1-k_{i-1})} \cdot$$

$$\cdot \frac{1}{c-n\delta}\left(\frac{x}{c-n\delta}\right)^{k_{i-1}-1}\left(1 - \frac{x}{c-n\delta}\right)^{L-k_{i-1}} dx,$$

$$v = \min\{\ 1 - M - (i-1)\delta - (n-i)\gamma\ ,\quad c-n\delta\},$$

where $f_i(z)$ is the probability density function of the random variable ζ_i. The probability in the second row of (4.6) can be expressed as an absolute probability and thus we obtain an expression similar to (3.4).

We remind that the random variables $\xi_1,\ldots\xi_n$ arise from a sample of size N, taken from a population uniformly distributed in the interval (0,1), in a way described in Section 2. The joint distribution of $\xi_1,\ldots,\xi_n$ given that $\xi_1+\ldots+\xi_i = u = M+(i-1)\delta-i\gamma+x$ coincide with the joint distribution of two independent random vectors. These vectors consists of j_i-1 resp. $N-j_i$ components and in both cases the joint densities are given by expressions of the type (2.6). In case of the first vector $N,n,1-n\gamma$ should be replaced by j_i-1, $i-1,u$ and in case of the second vector, by $N-j_i$, $n-i$, $1-n\gamma-u$ respectively. Similar is the situation concerning the random variables $\eta_1,\ldots,\eta_{n-1}$

We apply simulation for the computation of the probability $h(\underline{M})$. The computation of the gradient values is more sophisticated because beyond simulation numerical integration is also needed. Hence it seems to be more economic to apply gradient free minimization of the penalty function.

5. SIMULATION TECHNIQUE FOR THE COMPUTATION OF THE VALUES OF THE FUNCTION h(M).

Two methods are proposed. The first one follows the modelling of the delivery processes. We take many samples of size N resp. L, order them and select the required elements. This method has the great disadvantage that the ordering of the sample elements requires much computer time. It is known that the ordering time of N elements increases in the order of magnitude of $N\cdot\log_2 N$.

The second method is more effective than the just mentioned former one. It is based on the fact that any Dirichlet distribution can be represented as the joint distribution of random variables $y_1,\ldots,y_n$ by

$$y_i = x_i \,/\, (x_1 +\ldots+ x_{n+1}) \quad , \quad i = 1,2,\ldots,n \tag{5.1}$$

where $x_1,\ldots, x_{n+1}$ are independent, standard gamma distributed random variables with parameters $\nu_1 > 0,\ldots, \nu_{n+1} > 0$ i.e. x_i has the

following probability density:

$$(5.2) \qquad (z_i^{\nu_i-1} e^{-z}) \; / \; \Gamma(\nu_i)$$

In fact, the joint density of the random variables (5.1) is given by

$$(5.3) \qquad \frac{\Gamma(\nu_1+\ldots+\nu_{n+1})}{\Gamma(\nu_1)\ldots\Gamma(\nu_{n+1})} \; z_1^{\nu_1-1} \ldots z_n^{\nu_{n-1}} (1-z_1-\ldots-z_n)^{\nu_{n+1}-1}$$

if $z_i > 0$, $i = 1,\ldots,n$, $z_1+\ldots+z_n < 1$ and is 0 otherwise. Thus by a suitable choice of $\nu_1,\ldots,\nu_{n+1}$ the required Dirichlet density turns out.

Ahrens and Dieter [1] gave effective simulaiton technique for the simulation of the gamma distribution. Their method is particularly effective when the ν parameter is large or is not an integer.

The probability density functions, (2.5) and (2.6) slightly differ from the density function (5.3). The simulation technique described above requires only very simple modification in both cases. Let us consider the gamma probability density function

$$(5.4) \qquad (\lambda^{\nu} z^{\nu-1} e^{-\lambda z}) \; / \; \Gamma(\nu) \quad , \quad z > 0.$$

If $x_1,\ldots,x_{n+1}$ are independent and gamma distributed random variables with parameter pairs $\lambda,\nu_1;\ldots;\lambda,\nu_{n+1}$; where $\lambda = 1-n\gamma$, $\nu_1 = j_1$, $\nu_2 = j_2-j_1$,..., $\nu_n = j_n-j_{n-1}$, $\nu_{n+1} = N-j_{n+1}+1$ then the random variables defined by (5.1) have the same joint probability distribution as $\xi_1,\ldots,\xi_n$ do. On the other hand x_i can be represented as the sum of ν_i independent and exponentially distributed random variables with the same parameter λ , for every $i=1,\ldots,n+1$. Finally the exponentially distributed random variables can be represented as negative logarithms of random variables uniformly distributed in the interval (0,1). The simulation of the joint distribution of the random variables can be carried out in a similar way.

In case of this second simulation technique we only take the logarithms of the N+1 sample elements but do not order them. The required computer time is much less than in the first case.

The probabilities are approximated by relative frequencies. The sample size ensuring prescribed precision can be determined by the inequality of Bernstein.This is the following. If ν_m denotes the frequency of an event of probability p, in the course of m independent experiments and ε is a given positive number, then

$$(5.5)\quad P\left(\left|\frac{\nu_m}{m} - p\right| \geq \varepsilon\right) \leq 2\exp\left[-\frac{m\varepsilon^2}{2p(i-p)\left(1+\varepsilon/(2p(1-p))\right)^2}\right],$$

for $0 < \varepsilon \leq p(1-p)$.

If the probability on the left hand side equals δ, then for m we obtain the inequality

$$(5.6)\qquad m \geq 1/\varepsilon^2 - 2p(1-p)(1+\varepsilon/[2p(1-p)])^2 \cdot \log 2/\delta$$

if $0 < \varepsilon \leq p(1-p)$. For fixed δ and ε the largest value of the right hand side of (5.6) corresponds to $p = 1/2$, and it is a monotonically decresing function of p for $1/2 < p < 1$ (and monotonically increasing for $0 < p < 1/2$), provided $\varepsilon \leq p(i-p)$. In such a way we can get a lower bound for m which is good for every p. This is important because our aim is to approximate the probability p. Sometimes we have certain bounds for p. This is the case in connection with such stochastic programming problems where we have probabilistic constraint i.e. lower bound for the probability.

In our models we use at least, 0.8 as lower bound for the function $h(\underline{M})$. In practice this means that the factors are greater than or equal to 0.9. Using this information, the required sample size is much smaller then would be the case without any previous information. The table given below well illustrates the variation of the lower bound for m as a function of ε and p when δ is fixed at 0.1.

p	0.09	0.045	0.025	0.01
0.5	258	879	2 645	15 606
0.8	195	616	1 783	10 209
0.9	150	417	1 120	6 016
0.95	-	306	727	3 481

6. NUMERICAL EXAMPLE. As an example we consider a certain kind of product for the production of which two basic materials are needed and we want to determine the initial stock levels of the two basic materials ensuring the continuous production. The shortage in each of them stops the production and the cost of such an event is relatively high so that one of our main objectives is to avoid shortage by a prescribed probability near unity.

We assume that the demands for both basic materials are uniform in time and the unsatisfied demand remains i.e. the production plan has to be fulfilled. We assume that the basic materials are delivered from two different sources so that the two delivery processes can be supposed stochastically independent. The (0,1) time interval is now a quarter of a year, 90 days in other terms. According to long term statistics deliveries occur 4 resp. 5 times concerning the first resp. second basic material during one period (90 day). The table below shows actual delivery days for six past periods concerning the first basic material

Deliveries

Periods	1	2	3	4
1	23	41	61	82
2	27	48	73	88
3	30	39	60	90
4	19	48	68	89
5	24	50	65	78
6	28	42	71	82
Column averages	25,17	44,66	66,33	84,83

The minimum distance between two cunsecutive deliveries is 9 days. Since 90 days form a time interval of lenght 1, this means that the mentioned minimum distance is $\gamma^{(1)} = 0{,}1$. For the average delivery times we get in the same way

$$z_1 = 0.28, \qquad z_2 = 0.49, \qquad z_3 = 0.73, \qquad z_4 = 0.94.$$

Using our modeling of the delivery time process we can write

$$\bar{z}_i = i\gamma^{(1)} + E(x^*_{j_i}) \quad , \quad i = 1,\ldots,n,$$

where $x^*_{j_i}$ denotes the j_i th element of the ordered sample of size N taken from the population uniformly distributed in the interval $(0{,}1 - n\gamma^{(1)})$. We have to find integers $N, j_1,\ldots,j_n$ for which the following equalities hold at least in good approximation:

$$E(x^*_{j_i}) = j_i(1-n\gamma^{(1)})/(N+1) = \bar{z}_i - i\gamma^{(1)} \quad ,$$

$$j_i = \frac{\bar{z}_i - i\gamma^{(1)}}{1-n\gamma^{(1)}}(N+1) \ , \quad i = 1,\ldots,n.$$

Since the $\bar{z}_i$, $i = 1,\ldots,n$ and $\gamma^{(1)}$ are rationals in parctice such integers N, $j_1,\ldots,j_n$ always exist. It is not worth always to require that the above equalities hold exactly. In fact, if we work with large numbers, then the computer time will considerably be increased. In the above example the values of

$$(\bar{z}_i - i\gamma^{(1)} / (1-n\gamma^{(1)}) \ , \quad i = 1,2,3,4$$

are 0.298; 0.493; 0.728; 0.903 and choosing N=9, $j_1=3$, $j_2=5$, $j_3=7$, $j_4=9$, the above equalities are well approximated.
The next table shows the delivered amounts of the first basic material in the same past six periods

Deliveries

Periods	1.	2.	3.	4.	Totals
1	630	400	670	800	2500
2	700	500	600	900	2700
3	730	580	550	740	2600
4	720	620	650	1010	3000
5	760	580	760	1100	3200
6	750	650	780	920	3100

Dividing the rows by the sums of the rows we get the table

0.252	0.16	0.268	0.32
0.259	0.185	0.222	0.333
0.28	0.223	0.211	0.248
0.24	0.206	0.216	0.336
0.237	0.181	0.238	0.344
0.242	0.21	0.252	0.297
Column averages 0.252	0.194	0.234	0.319

From the table we see that the minimal delivared amount is $\delta^{(1)} = 0.16$.

If the column averages are denoted by $\bar{u}_1, \ldots, \bar{u}_n$ and we introduce the further motation $v_i = \bar{u}_1 + \ldots + \bar{u}_i$, $i = 1, \ldots, n$, then similarly to the case of the delivery times write the equalities

$$v_i = i\delta^{(1)} + E(y^*_{k_i}), \quad i = 1, \ldots, n-1$$

where $y^*_{k_i}$ denotes the k_i th element of a sample of size L taken from a population uniformly distributed in the interval $(0, 1 - n\delta^{(1)})$. We want to determine integers $L, k_1, \ldots, k_{n-1}$ so that the following equalities hold at least in good approximation. In our case the values of

$$(v_i - i\,\delta^{(1)}) \;/\; (1 - n\delta^{(1)}), \quad i = 1,2,3$$

are 0.255; 0.35; 0.555. Thus the choices $L = 20$, $k_1 = 5$, $k_2 = 7$, $k_3 = 11$ provide good approximations.

We can proceed in a similar way concerning the second basic material. Assume that we obtained the following values: $n = 5$, $N = 10$, $j_1 = 2$, $j_2 = 3$, $j_3 = 5$, $j_4 = 7$, $j_5 = 9$, $\gamma^{(2)} = 0.15$. $L = 10$, $k_1 = 2$, $k_3 = 7$, $k_4 = 8$, $\delta^{(2)} = 0.12$. As regards the objective function, we have choosen $d^{(2)} = 3d^{(1)}$.

The SUMT interior point method started with values between 0.6 and 0.8. The $r_1, r_2, \ldots$ sequence was choosen to be 1, 1/5, 1/25,... and the initial values of the k+1st unconstrained minimization were the optimal values of the kth unconstrained minimization. The method was stopped when the change in the optimal values of the penalty function was less than 0.01. The method of Hooke and Jeeves [3] was applied for the minimization of the penalty function. In our numerical example the minimizing $M^{(1)}$ and $M^{(2)}$ belonging to $r = 1/125$ were accepted as optimal solutions of the problem. These are $M^{(1)} = 0.32$ and $M^{(2)} = 0.19$. This means that 32% of the total demand of the first material and 19% of the total demand of de second material will serve the production without shortage with probability $p = 0.8$ and the cost will be minimum among all feasible alternatives.

The test programs written in FORTRAN run between 1.5 and 2.5 minutes on a CDC 3300 computer. Further unconstrained optimization methods were also tested such as the method of Rosenbrock [11] and Powell [5]. The best computer time was produced by the method of Hooke and Jeeves, however. This method was successfully applied also in other stochastic programming problems where function values were determined by simulation.

REFERENCES

[1] Ahrens, J.H. and Dieter, K. "Computer Methods for Sampling from Gamma, Beta, Poisson and Binomial Distributions", *Computing*, Vol. 12(1974), pp. 223-246.

[2] Fiacco, A.V. and McCormick, G.P. *Nonlinear Programming: Sequential Unconstrained Minimization Technique*, Wiley, New York, 1968.

[3] Hooke,R. and Jeeves, T.A. "Direcht Search Solution of Numerical and Statistical Problems", *Journal of the A.C. M.* Vol. 8(1959), pp. 215-229.

[4] Miller, B.L. and Wagner, H.M. "Chance Constrained Programming with Joint Constrainsts", *Operations Research*, Vol. 13(1965), pp. 930-945.

[5] Powell, M.J.D. "An Iterative Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivetives", *Computer Journal*, Vol. 7(1964), pp. 155-162.

[6] Prékopa, A. "Reliability Equation for an Inventory Problem and its Asymptotic Solution," *Coll. on Appl. of Math. to Economics*, Akadémiai Kiadó, Budapest, 1965. pp.317-327.

[7] Prékopa,A. "On Logarithmic Concave Measures with Application to Stochastic Programming", *Acta Universitatis Szegedienis*, Vol. 32(1971), pp. 301-316.

[8] Prékopa,A. "Contributions to the Theory of Stochastic Programming", *Mathematical Programming*, Vol. 4(1973), pp.202-221.

[9] Prékopa,A. "Stochastic Programming Models for Inventory Control and Water Storage Problems", *Colloquia Mathematica Societatis János Bolyai, 7. Inventory Control and Water Storage, Győr, 1971.* Bolyai Janos Math.Soc. and North Holland Publ. Comp., Budapest, 1973.

[10] Prékopa, A. "Generalizations of the Theorems of Smirnow with Applications to a Reliability Type Inventory Problem," *Mathematische Operations forschung und Statistik*, **Vol.** 4(1973) , pp.283-297.

[11] Rosenbrock, H.N. "An Automatic Method of Finding the Greatest or Least Value of a Function", *The Computer Journal*, Vol. 3(1960) , pp. 175-184.

[12] Wilks, S.S. *Mathematical Statistics*, Wiley, New York, 1962.

[13] Ziermann,M. "Anwendung des Smirnov'schen Sätzen auf Lagerhaltungsproblem", *Publications of the Math. Inst. of the HAS.*, Vol. 8(1965), Series B, pp.509-518.

A New Method for Serially Linked Reservoir System Design Using Stochastic Programming*

András Prékopa
Technological University of Budapest
Hungarian Academy of Sciences

Tamás Rapcsák
Hungarian Academy of Sciences

István Zsuffa
National Water Authority of Hungary

ABSTRACT

A stochastic programming model formulated for a serially linked reservoir system design is presented and a numerical solution method is proposed. The model was already formulated in [8] in a concise form. Here we give a more detailed explanation, apply SUMT for the solution of the optimization problem and a present a numerical exemple. The example is taken from a larger model which arose in the first implementation of this reservoir system design method in practice.

Key words: reservoir system design, stochastic programming, simulation, nonlinear programming.

1. Introduction. The reservoir system design model considered here was introduced in [8]. The purpose of the present paper is to explain this model in more detail, give a solution procedure for the nonlinear programming problem representing the model and present a numerical example of this new reservoir system design method.

* This work was supported by the Hungarian Academy of Sciences and the National Water Authority of Hungary

Reservoir system design and operation models have been treated in the literature prior to the first publication of the model we consider here. However, these other models contain probabilistic constraints of a very simple type, neglecting the stochastic dependence between random variables and using separate probabilistic constraints for every constraint containing random variables.

A reservoir design principle can be derived from the classic results of Moran [5]. In this case, however, we are working with stichastically independent and time homogeneous water inputs and even the various generalizations do not go very far from these assumptions.

In this paper our aim is to remove these strong conditions. We allow correlated and non time homogeneous inputs. Only a certain general property will be required from the joint probability density of random quantities, among which the demand variables may be included.

In the past few years the first named author obtained new mathamatical results in conncection with "logarithmic concave measures" and formulated stochastic programming models on the basis of this mathematical theory. This reservoir system design model is one of the applications of this new way of problem formulation and solution to specific water resources problems.

Teh model will be described in more detail for the case of serially linked reservoirs because the optimal operating policy is simple in this case under the assumptions formulated in the next section. We can handle problems, however, in case of a more sophisticated topological structure of rivers and sites if we assume the operating policy to be used.

The problem will be to find optimal reservoir capacities such that all demands should be met in the course of a prescribed number of periods with a probability which is greater than or equal to a prescribed (in practice high) probability. We assume that released water leaves the system thus we are considering reservoir used for irrigation,

municipal etc. purposes and exclude e.g. hydroelectric power generation. The objective function to be minimized is the sum of total building costs and the expectation of penalties occuring whenever demand is not completely satisfied.

In Section 2 we present the hypotheses and formulate the problem. In Section 3 we discuss the mathematical properties of the problem and present the techniques of the numerical solution. In Section 4 a numerical exemple is given and in the Appendix we summarize some basic mathematical notions and prove the theorems formulated in the text.

2. FORMULATION OF THE PROBLEM. The topology of the main river, the side rivers and the possible reservoir sites is illustrated on *Fig.1.*

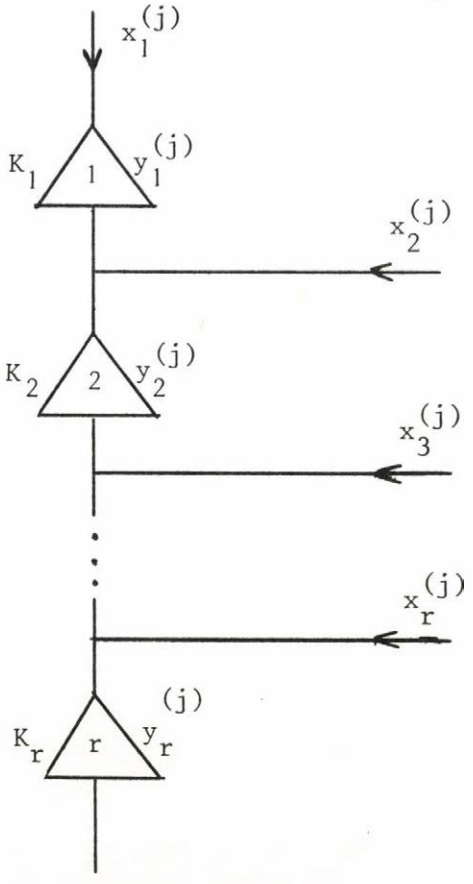


Figure 1.

Time is subdivided into periods and we consider a finite number of them. These can mean months, decades etc. in practice.

We assume that in the beginning of every period certain water inputs occur in accordance with the topology of the rivers and reservoirs. If a reservoir becomes full then additional water overflows to downstream reservoirs. No more water is released from upstream reservoirs to downstream reservoirs at these times.

At the end of every period demands occur which can be assigned to separate reservoirs. Every demand is satisfied from the assigned reservoir if it is possible. If not, then our assumed operating policy is as follows. First demands are satisfied to the extent of water amounts being in the corresponding reservoirs. Then starting from the reservoir furthest downstream, stop at the first reservoir where there is unsatisfied demand. From here we aggregate the unsatisfied demands of all consecutive reservoirs up to the first nonempty reservoir and try to meet this aggregated demand. If this is not possible then we proceed similarly in the upstream direction. If the whole system can meet the total demand then this procedure stops at a certain point and we can then satisfy the downstream demands. This procedure is repeated for the remaining upstream subsystem etc. In our model all demand will be met by a prescribed high probability. Note that the above operating policy is uniquely determined also in the case where part of the demand remains unsatisfied.
We assume further that if the system is unable to satisfy the demands in a certain period then a penalty occurs which belongs to the whole system and is a linear function of the unsatisfied part of the total demand. The proportionality factor may depend on the period in question. Nonlinear penalties can similarly be handled.

Let us introduce the following notations:

r	number of sites;
K_i	unknown capacity of Reservoir i;
V_i	prescribed constant, upper bound for K_i;
n	number of periods;

$z_i^{(o)}$ initial water content of Reservior i;

$z_i^{(j)}$ water content of Reservoir i at the end of the j-th period provided there is no unsatisfied demand in the course of the n periods; the z will take on negative values too, the physical interpretation for this case will be given in the text;

$x_i^{(j)}$ direct input into Reservoir i in the j-th period;

$y_i^{(j)}$ direct demand against Reservoir i in the j-th period;

$q^{(j)}$ penalty of one unit of unsatisfied demand in the j-th period;

$c_i(K_i)$ building cost of Reservoir i as a function of its capacity;

where $i = 1,\ldots,r$ and $j = 1,\ldots,n$.

The $x_i^{(j)}$, $y_i^{(j)}$ will be assumed to be random variables. The $z_i^{(j)}$ are functions of these (and of the $z_i^{(o)}$, K_i) hence they are also random variables. For a moment let us fix these random variables at some values. We express them recoursively in the following manner. First we define the quantities g_i, h_i by the relations

$$
\begin{aligned}
& g_o^{(j)} = 0, \\
& g_i^{(j)} = z_i^{(j-1)} + g_{i-1}^{(j)} + x_i^{(j)} - \min(z_i^{(j-1)} + g_{i-1}^{(j)} + x_1^{(j)}, K_i), \\
& h_i^{(j)} = \min(z_i^{(j-1)} + g_{i-1}^{(j)} + x_i^{(j)}, K_i), \\
& \qquad i = 1,\ldots, r \;;\; j = 1,\ldots, n.
\end{aligned}
\tag{2.1}
$$

The $g_i^{(j)}$ is the amount of water overflown at reserwoir i and the $h_i^{(j)}$ is the amount of water remaining in Reservoir i in the beginning of Period j when the input water fills up the reservoirs and demand does not occur yet. Next we define the quantities $d_i^{(j)}$ by

$$(2.2) \qquad d_i^{(j)} = h_i^{(j)} - y_i^{(j)} \quad , \quad i = 1 \,,\ldots, r \,; \quad j = 1 \,,\ldots, n.$$

Using these, finally we write

$$(2.3) \qquad \begin{aligned} z_1^{(j)} &= \min(d_1^{(j)} \,,\; d_1^{(j)}+d_2^{(j)} \,,\ldots,\; d_1^{(j)} +\ldots+ d_r^{(j)} \,), \\ z_i^{(j)} &= \max[0, \min(d_i^{(j)}, d_i^{(j)}+d_{i+1}^{(j)} \,,\ldots,\; d_i^{(j)} +\ldots+ d_r^{(j)})], \\ & \qquad i = 2, \ldots, r \;\; ; \;\; j = 1 \,,\ldots, n. \end{aligned}$$

The above physical interpretation for the quantities $g_i^{(j)}$, $h_i^{(j)}$ is correct only in the case when it happens that $z_1^{(j)} \geq 0$ for $j=1,\ldots,n$. In this case $z_i^{(j)}$ equals the water content of Reservoir i at the end of period j for $i=1,\ldots,r$ and $j=1,\ldots,n$.

If $z_1^{(j)} < 0$ then $-z_1^{(j)}$ means the accumulated total unsatisfied demand with respect to the whole system up to the end of period. In the case when this happens for at least one $j(1 \leq j \leq n)$, then we assign those values to $g_i^{(j)}$, $h_i^{(j)}$, $d_i^{(j)}$, $z_i^{(j)}$ which come out from the recursive relations in a formal mathematical manner, but only $z_1^{(j)}$ is interesting for us from the point of view of physical interpretation. This will enter the objective function.

All demands will be met in the course of the n periods if and only if

$$(2.4) \qquad z_1^{(j)} = \min(d_1^{(j)}, d_1^{(j)}+d_2^{(j)} \,,\ldots,\; d_1^{(j)}+\ldots+d_r^{(j)}) \geq 0 \quad , \quad j=1,\ldots,n$$

which is equivalent to the system of inequalities

$$(2.5) \qquad d_1^{(j)} \geq 0 \;, \;\; d_1^{(j)}+d_2^{(j)} \geq 0, \;\ldots,\; d_1^{(j)}+\ldots+d_n^{(j)} \geq 0, \quad j=1,\ldots,n.$$

In the above relations only the $z_i^{(o)}$, $x_i^{(j)}$, $y_i^{(j)}$, K_i are independent variables, all others are expressed in terms of these. The number of independent variables is $2r+2rn$. The $d_i^{(j)}$, $z_i^{(j)}$, $g_i^{(j)}$, $h_i^{(j)}$ will be considered as functions of $2r+2rn$ variables though not

all variables appear in their expressions. The variables $K_1,\ldots,K_r$ will be restricted to be nonnegative where as the others are not restricted in the formulas (2.1)-(2.3). They will be restricted later on by the constraints of our problem.

We are now in the position that we can formulate our reservoir system design problem. This is the following nonlinear programming problem:

$$\text{minimize} \quad \Big[\sum_{i=1}^{r} c_i(K_i) + \sum_{j=1}^{n} q^{(j)} E(\mu^{(j)}) \Big]$$

subject to the constraints

$$(2.6) \qquad \begin{aligned} &P(d_1^{(j)} \geq 0,\ d_1^{(j)}+d_2^{(j)} \geq 0,\ldots,\ d_1^{(j)}+\ldots+d_r^{(j)} \geq 0,\ \ j=1,\ldots,n) \geq q, \\ &0 \leq K_i \leq V_i \ , \quad i=1,\ldots,r, \end{aligned}$$

where q is a prescribed probability level $(0<q<1)$ near unity in practice, $q^{(1)},\ldots,q^{(n)}$ are some nonnegative constants, (they can be discounting factors if the total length of the periods is long enough), E is the symbol of expectation and

$$(2.7) \qquad \mu^{(j)} = \begin{cases} -z_1^{(j)} & \text{if} \quad z_1^{(j)} < 0, \\ 0 & \text{otherwise,} \end{cases} \qquad j = 1,\ldots,n,$$

We assume that $c_i(K_i)$ is a continuous function of K_i for $K_i \geq 0$, $i = 1,\ldots,r$.

3. MATHEMATICAL PROPERTIES OF THE RESERVOIR SYSTEM DESIGN MODEL (2.6) AND ITS ALGORITHMIC SOLUTION. Problem (2.6) is a nonlinear programming problem, where the objective function and one of the constraining functions are nonlinear. Particular attention has to be paid to the nonlinear constraining function the values of which are probabilities in higher dimensional space. These probabilities depend on the unknown capacities $K_1,\ldots,K_r$ which are the decision variables in Problem (2.6).

Two principal reasons require this particular attention:

a) The numerical solution of Problem (2.6) is simpler if the problem is a convex programming problem (minimization of a convex function on a convex set) and since the objective function will be convex under a wide arange of assumptions, it remains to investigate the constraining function standing on the left hand side in the probabilistic constraint, as two when it allows our problem to be convex.

b) Any optimization procedure we apply for the numerical evaluation of Problem (2.6) requires at least the function values of the probabilistic constraining function at every step in the optimization. This highly nontrivial problem is solved by Monte Carlo technique or in other term by simulation.

The mathematical properties of the probabilistic constraining function will be derived from the following two theorems.

Theorem 1. Let $h_1(\underline{u},\underline{v})$,..., $h_r(\underline{u},\underline{v})$ be functions of two vector variables: u,v, and suppose that these functions are concave with respect to all variables. Let $\underline{\xi}$ be a vector valued random variable having the same number of components as $\underline{v}$ and consider the probability

$$P(g_1(\underline{u},\underline{\xi}) \geqq 0 \ ,\ldots,\ g_r(\underline{u},\underline{\xi}) \geqq 0) \tag{3.1}$$

as a function of $\underline{u}$.

If the components of $\underline{\xi}$ have a continuous joint probability distribution and the joint probability density is a logarithmic concave function (see the Appendix) then the probability (3.1) is a logarithmic concave function of the vector variable $\underline{u}$.

This theorem was essentially proved in [6]. It was first formulated exactly in this form in [7]. For further similar results the reader is referred to [9].

Theorem 2. The functions $d_1^{(j)}$, $d_1^{(j)} + d_2^{(j)}$,..., $d_1^{(j)} + \ldots + d_r^{(j)}$, $j = 1, \ldots, n$ are concave functions of the $2r+2rn$ variables.

This theorem will be proved in the Appendix. Theorem 1 and Theorem 2 together imply.

Theorem 3. If the joint probability distribution of the random variables $x_1^{(j)}, \ldots, x_r^{(j)}, y_1^{(j)}, \ldots, y_r^{(j)}$ is continuous and their joint probability density is a logconcave function, then the probability

$$(3.2) \qquad P(d_1^{(j)} \geq 0,\ d_1^{(j)}+d_2^{(j)} \geq 0, \ldots, d_1^{(j)}+\ldots+d_r^{(j)} \geq 0,\ j=1,\ldots,n)$$

is a logconcave function of the variables $z_1^{(o)}, \ldots, z_r^{(o)}, K_1, \ldots, K_r$.

Remark. Theorem 3 implies that if we set $z_1^{(o)} = K_1, \ldots, z_r^{(o)} = K_r$ which means that we start with full reservoirs and consider the probability (3.2) with this assumption, then this will be a logarithmic concave function of the variables $K_1, \ldots, K_r$.

In [6] a number of multivariable probability densities are mentioned which are logarithmic concave functions. Here we shall use the multivariate normal distribution, the density of which is given by

$$(3.3) \qquad f(\underline{u}) = \frac{\sqrt{|C|}}{(2\pi)^{k/2}} e^{-\frac{1}{2}(\underline{u}-\underline{\mu})' C^{-1} (\underline{u}-\underline{\mu})}, \quad \underline{u} \in R^k,$$

where C is the covariance matrix and $\underline{\mu}$ is the expectation vector. It is well-known that if C is a positive definite matrix, then the same holds for C^{-1} and it is also well-known that a quadratic form is a convex function if its matrix is positive definite. These imply that $\ln f(\underline{u})$ is concave in the entire k-dimensional space.

In connection with the objective function of Problem (2.6) we mention the following theorem the proof of which is given in the Appendix.

Theorem 4. No matter what kind of joint probability distribution the random variables $x_i^{(j)}$, $y_i^{(j)}$, $i=1,\dots,r$ Q $j=1,\dots,n$ have, the function

$$\sum_{j=1}^{n} q^{(j)} E(\mu^{(j)})$$

is a convex function of the variables $z_i^{(o)}$, K_i , $i = 1,\dots,r$, provided $q^{(j)} \geqq 0$, $j = 1,\dots,n$.

For the solution of our problem we apply the Sequential Unconstrained Minimization Technique (SUMT) with logarithmic penalty function. For the general explanation of this technique see [2]. Here we are interested only in the solution of Problem (*2.6*).

First we reduce to zero the constraints of Problem (*2.6*) then form the following function

$$\sum_{i=1}^{r} c_i(K_i) + \sum_{j=1}^{n} q^{(j)} E(\mu^{(j)}) -$$

$$(3.4) \qquad - t\{\ell n[P(d_1^{(j)} \geqq 0,\dots,d_1^{(j)} + \dots + d_n^{(j)} \geqq 0, j=1,\dots,n) - q] +$$

$$+ \sum_{i=1}^{r} \ell n\, K_i + \sum_{i=1}^{r} \ell n(V_i - K_i)\},$$

where t is a fixed positive number. This has an unconstrained minimum which is attained at some capacity vector, depending on t, $\underline{K}(t)$, say. We take a sequence of t values, $t_1 > t_2 > t_3 > \dots$ for which $t_p \to 0$ if $p \to \infty$ and minimize the function (*3.4*) sequentially for every t_p. Then as p increases we approach the optimum value of the original Problem (*2.6*). More exactly we have

Theorem 5. When replacing $K_i = V_i$, $i=1,\dots,r$ in Function (*3.2*), we obtain a probability strictly greater then q, then the sequence of values of Function (*3.4*) in case of $t_1, t_2, \dots$ tends to the optimum value of Problem (*2.6*).

The practical content of Theorem 5 is that for some large p, the vector $K(t_p)$ can be regarded as optimal solution to Problem (*2.6*).

For every t_p we have to perform an unconstrained minimization of the function (*3.4*). Under the assumption of Theorem 3 this function is convex provided $c_1(K_1)$,..., $c_r(K_r)$ are convex functions of the variables $K_1,\ldots,K_r$. In fact the function (*3.2*) is logconcave hence (as it is easy to see)

$$(3.5)\qquad P(d_1^{(j)} \geqq 0,\ \ldots,\ d_1^{(j)} + \ldots + d_r^{(j)} \geqq 0,\quad j = 1,\ldots,n) - q$$

is also logconcave on the set where (*3.5*) is nonnegative. Thus the additional term that we add to our original objective function in (*3.4*), i.e., the sum

$$\sum_{j=1}^{n} q^{(j)}\ E\ (\mu^{(j)})$$

is a convex function of the variables $z_i^{(o)}$, K_i, $i=1,\ldots,r$.

Thus when performing the SUMT, at every step we have to minimize (unconstrained) a convex function. This fact is very pleasent from the numerical point of view. Since it is not realistic to compute gradients in our case, we have to apply a minimization method using only function values. For this purpose we use the method of Hooke and Jeaves (see [3]). The detailed description of finding values of the function (*3.2*) i.e., probabilities of sets in higher dimensional spaces in case of multivariate normal distribution, can be found in [1]. In [10] a multivariate gamma distribution and its fitting to empirical data is proposed. The use of this is also advised.

4. NUMERICAL EXAMPLE. The following example arose from the first implementation work (see *Figure 2.*) of the reservoir system design model described in the previous sections. We have three periods, June, July, August and twelwe random variables which will be denoted here by x_1^{jun}, y_1^{jun}, x_2^{jun}, y_2^{jun}, x_1^{jul}, y_1^{jul}, x_2^{jul}, y_2^{jul}, x_1^{aug}, y_1^{aug}, x_2^{aug}, y_2^{aug}. We assume that $z_1^{(o)} = K_1$, $z_2^{(o)} = K_2$. We assume further that our twelwe variables have a joint normal distribution with the following expectations, variances and correlation matrix:

	Expectations m^3	Standard Deviations m^3
x_1^{jun}	464822	186984
y_1^{jun}	215760	327120
x_2^{jun}	929644	373960
y_2^{jun}	152033	275890
x_1^{jul}	320576	266040
y_1^{jul}	433608	243600
x_2^{jul}	641152	532080
y_2^{jul}	396225	205450
x_1^{aug}	266040	234040
y_1^{aug}	484416	214368
x_2^{aug}	532080	511060
y_2^{aug}	407965	180796

$$
R = \begin{bmatrix}
1,00 & 0,10 & 0,80 & 0,05 & 0,60 & 0,12 & 0,50 & 0,06 & 0,40 & 0,06 & 0,30 & 0,03 \\
 & 1,00 & 0,05 & 0,80 & 0,12 & 0,25 & 0,10 & 0,23 & 0,08 & 0,02 & 0,05 & 0,00 \\
 & & 1,00 & 0,13 & 0,55 & 0,15 & 0,68 & 0,13 & 0,50 & 0,00 & 0,52 & 0,00 \\
 & & & 1,00 & 0,15 & 0,20 & 0,13 & 0,18 & 0,06 & 0,00 & 0,06 & 0,00 \\
 & & & & 1,00 & 0,10 & 0,80 & 0,09 & 0,70 & 0,00 & 0,65 & 0,00 \\
 & & & & & 1,00 & 0,09 & 0,70 & 0,15 & 0,20 & 0,13 & 0,02 \\
 & & & & & & 1,00 & 0,10 & 0,65 & 0,00 & 0,70 & 0,00 \\
 & & & & & & & 1,00 & 0,13 & 0,18 & 0,10 & 0,20 \\
 & & & & & & & & 1,00 & 0,10 & 0,80 & 0,08 \\
 & & & & & & & & & 1,00 & 0,10 & 0,85 \\
 & & & & & & & & & & 1,00 & 0,10 \\
 & & & & & & & & & & & 1,00
\end{bmatrix}
$$

where the ordering of the random variables is the same as before.

We have furthermore

$$V_1 = 1500000 \text{ m}^3$$

$$V_2 = 2500000 \text{ m}^3$$

$$c_1(K_1) = \begin{cases} K_1 & \text{if } 0 \leqq K_1 \leqq 500000 , \\ 500000 + 0,4(K_1 - 500000) & \text{if } K_1 > 500000, \end{cases}$$

$$c_2(K_2) = \begin{cases} 0,45K_2 & \text{if } 0 \leqq K_2 \leqq 1000000, \\ 450000 + 0,6(K_2 - 1000000) & \text{if } 1000000 < K_2 \leqq 1500000 \\ 750000 + 0,8(K_2 - 1500000) & \text{if } K_2 > 1500000, \end{cases}$$

where the unit of the cost is Ft100. (see *Figure 2* and *3*) Finally $z_1^{(o)} = K_1$, $z_2^{(o)} = K_2$ and $q=0,8$. Our problem is the following:

minimize $(c_1(K_1) + c_2(K_2))$
subject to the constraints

$$(4.1) \qquad P\left(\begin{array}{ll} d_1^{(1)} \geqq 0, & d_1^{(1)} + d_2^{(1)} \geqq 0 \\ d_1^{(2)} \geqq 0, & d_1^{(2)} + d_2^{(2)} \geqq 0 \\ d_1^{(3)} \geqq 0, & d_1^{(3)} + d_2^{(3)} \geqq 0 \end{array} \right) \geqq 0,8 ,$$

$$0 \leqq K_1 \leqq V_1 ,$$

$$0 \leqq K_2 \leqq V_2 .$$

Note that $c_1(K_1)$ is not a convex function. However, the values $0 \leq K_1 \leq 0{,}5 \cdot 10^6$ are automatically discarded since the probabilistic constraint is not satisfied even if we choose $K_2 = V_2$. Thus we may restrict ourselves to the halfline $K_1 \geq 0{,}5 \cdot 10^6$. In this case $c_1(K_1)$ is linear hence also convex. The unconstrained minimization procedures have to be carried out for the funciton

$$(4.2) \qquad c_1(K_1)+c_2(K_2)-t\{\ell n(P-0{,}8)+\ell nK_1+\ell nK_1+\ell nK_2 + \\ + \ell n(1{,}5 \cdot 10^6-K_1)+\ell n(2{,}5 \cdot 10^6-K_2)\}$$

in cases of given t values forming a decreasing sequence and approaching zero.

Three unconstrained optimization had to be performed. The corresponding t values were $t_1=1$, $t_2=1/5$, $t_3=1/25$.
The initial solution was

$$K_1 = 1{,}4 \cdot 10^6 \; ; \quad K_2 = 2{,}4 \cdot 10^6.$$

The corresponding value of the probabilistic constraining function is 0,984. The second (third) unconstrained optimization started with the optimal solution of the first (second) optimization.

The Hooke-Jeaves method performed 7,2,1 iterations, respectively in the three unconstrained problems. The stopping rule was formulated in terms of changes in the variables in the consecutive steps.

As the optimal solution we received the following values

$$K_1 \text{opt} = 1{,}046289 \cdot 10^6 \text{m}^3$$

$$K_2 \text{opt} = 0{,}611206 \cdot 10^6 \text{m}^3$$

At these values the probabilistic constraint holds with equality sign i.e. 0,8 is the reliability level of these reservoir capacities. The total building cost is Ft $99{,}3556 \cdot 10^6$.

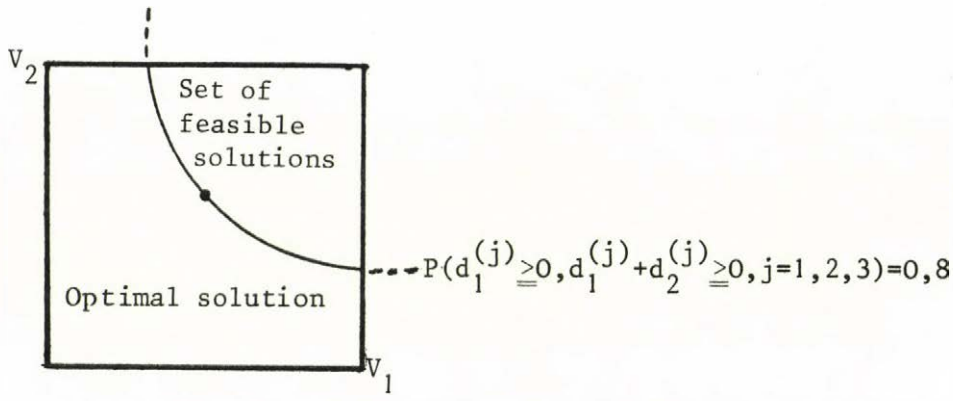


Figure 4.

Illustration of the set of feasible solutions of Problem (*4.1*).

APPENDIX

BASIC NOTIONS AND PROOFS OF THEOREMS

We denote by R^k the k-dimensional Euclidean space of the ordered k-tuples of real numbers.

A subset A of R^k is said to be convex if for every pair $\underline{u}_1, \underline{u}_2 \in R^k$ and $0<\lambda<1$, we have $\lambda\underline{u}_1 + (1-\lambda)\underline{u}_2 \in R^k$.

A function f defined on a convex set $A \subset R^k$ is said to be convex (concave) if for every pair $\underline{u}_1, \underline{u}_2 \in A$ and $0<\lambda<1$, we have

$$f(\lambda\underline{u}_1+(1-\lambda)\underline{u}_2) \underset{(\geq)}{\leq} \lambda f(\underline{u}_1)+(1-\lambda)f(\underline{u}_2).$$

If $f \geq 0$ and

$$f(\lambda \underline{u}_1+(1-\lambda)\underline{u}_2) \geq [f(\underline{u}_1)]^{\lambda}[f(\underline{u}_2)]^{1-\lambda},$$

then f is said to be logarithmic concave (or briefly logconcave) on A. It is easy to show that the set of those $\underline{u}$ vectors for which $f(\underline{u}) > 0$ is a convex subset of A.

Lemma. If a function f is logconcave on $A \subset R^k$ and for some $\underline{u}_1, \underline{u}_2$ we have $f(\underline{u}_1)>q$, $f(\underline{u}_2)=q$ where $0<q<1$ then there is no $\underline{u} \neq \underline{u}_2$ on the line segment joining $\underline{u}_1$ and $\underline{u}_2$ for which $f(\underline{u}_2)=q$.

Proof. Let $\underline{u} = \lambda \underline{u}_1 + (1-\lambda)\underline{u}_2$, $0 < \lambda < 1$.

Then we have

$$f(\underline{u}) \geq [f(\underline{u}_1)]^{\lambda}[f(\underline{u}_2)]^{1-\lambda} > q,$$

which proves the assertion.

Proof of Theorem 2. Before starting the proof we mention that the minimum of two concave functions is also concave.

The following equality systems can be verified easily by proceeding successively with p

$$h_1^{(j)}+\ldots+ h_p^{(j)} = \min(z_1^{(j-1)}+\ldots+z_p^{(j-1)}+x_1^{(j)} +\ldots+$$

$$(1) \qquad + x_p^{(j)}, K_p+h_1^{(j)} +\ldots+ h_{p-1}^{(j)}),$$

$$p = 2,\ldots,r \ ; \ j = 1,\ldots,n,$$

from where we derive

$$d_1^{(j)} +\ldots+ d_p^{(j)} =$$

$$= \min (z_1^{(j-1)} +\ldots+ z_p^{(j-1)} + x_1^{(j)} +\ldots+ x_p^{(j)} -$$

$$(2) \qquad - y_1^{(j)} - \ldots - y_p^{(j)} , K_p - y_p^{(j)} + d_1^{(j)}+\ldots+ d_{p-1}^{(j)}),$$

$$p = 2,\ldots, r \ ; \ j = 1, \ldots, n,$$

furthermore

$$z_1^{(j)} + \ldots + z_p^{(j)} =$$

$$= \min(d_1^{(j)} + \ldots + d_p^{(j)}, \ldots, d_1^{(j)} + \ldots + d_r^{(j)}), \tag{3}$$

$$p = 1, \ldots, r \; ; \; j = 1, \ldots, n.$$

In (1) the case $h_1^{(j)}$ and in (2) the case $d_1^{(j)}$ are not included but they have very simple expressions given by (2.1) , (2.2). It follows from (2.1) and (2.2) that $d_1^{(j)}$ is concave for $j = 1, \ldots, n$. Using this statement for the special case $j = 1$ and the equalities (2) we derive successively that the functions

$$d_1^{(1)} + \ldots + d_p^{(1)} \quad , \quad p = 1, \ldots, r$$

are concave.

Now we prove the assertion by induction with respect to j. We just proved the assertion for the case $j = 1$. Assume that the functions

$$d_1^{(j-1)} + \ldots + d_p^{(j-1)} \quad , \quad p = 1, \ldots, r$$

are concave. Then (3) implies that the functions

$$z_1^{(j-1)} + \ldots + z_p^{(j-1)} \quad , \quad p = 1, \ldots, r$$

are concave. Using the fact that $d_1^{(j)}$ is concave and the equalities (2), we derive successively that also the functions

$$d_1^{(j)} + \ldots + d_p^{(j)} \quad , \quad p = 2, \ldots, r$$

are concave. Thus the assertion holds for j too and the theorem is proved.

Proof of Theorem 4. The equality (3) and Theorem 2 imply that the functions

(4) $$-z_1^{(j)} \quad , \quad j = 1,\ldots,n$$

are convex. This implies that the functions $\mu^{(j)}$, $j = 1,\ldots,n$ defined by (2.7) are also convex.

In general if we replace by random variables some of the variables of a convex function and take its expectation, then this expectation is a convex funciton of the remaining variables. This proves the theorem.

Proof of Theorem 5. Proofs under general assumptions for the convergence of the SUMT are given in [2]. Our problem (2.6) is of the following type

$$\text{minimize } f(\underline{K})$$

(5) $$\text{subject to the constraints}$$

$$h_i(\underline{K}) \geqq 0 \quad , \quad i = 1,\ldots,m,$$

where $f, h_1, \ldots, h_m$ are defined and continuous for every $\underline{K} \in R^r$. Also the set of feasible solutions

(6) $$F = \{\underline{K} \mid h_i(\underline{K}) \geqq 0 \quad , \quad i = 1,\ldots,m \}$$

is bounded hence an optimal solution exists.

The penalty function $V_t(\underline{K})$ has the following form

(7) $$V_t(\underline{K}) = f(\underline{K}) - t \sum_{i=1}^{m} \ln h_i(\underline{K}),$$

where t is a positive constant. Let $\underline{K}_t$ denote a vector which minimizes the function (7).

Now it is proved in [2] that if t_p is a decreasing sequence of positive numbers for which $t_p \to 0$ if $p \to \infty$, then the sequence $V_{t_p}(\underline{K}_{t_p})$ tends to the optimum value of Problem 5 provided all interior points of the set (6)

are contained in the set

$$(8) \qquad M = \{ \underline{K} \mid h_i(\underline{K}) > 0 \quad , \quad i = 1,\ldots,m \}$$

which is supposed to be nonempty. Thus we only have to check this last property in case of Problem (2.6).

According to the assumption mentined in Theorem 5, the function (3.2) is strictly greater then p incase of $\underline{K} = \underline{V}$. Since the funciton (3.2) in continuous this implies that there exists a $\underline{K}_1$ for which all constraints of Problem (2.6) are satisfied by strict inequalities, thus the set (8) is nonempty in our case.

Now if $\underline{K}$ is an interior point of the set F in case of the special Problem (2.6), then trivially $0<K_i<V_i$, $i = 1,\ldots,r$ and by the Lemma the function (3.2) is also strictly greater than p in case of this $\underline{K}$. Thus the theorem is proved.

REFERENCES

[1] I. Deák, A Fast Monte Carlo Method for the Computation of Probabilities of Sets in Higher Dimensional Spaces in Case of a Multivariate Normal Distribution, Water Resources Research (to appear)

[2] A.V.Fiacco and G.P. McCormic. Nonlinear Programming. Sequential Unconstrained Minimization Techniques, Wiley, New York, 1968.

[3] J.Kowalik and M.R. Osborne, Methods for Unconstrained Optimization Problems, American Elsevier Publishing Co. Inc., New York, 1968.

[4] D.P. Loucks, Stochastic Methods for Analyzing River Basin Systems, Research Project Technical Completion Report, OWRR Project, No. C-1034 Dept. of Water Resources Engineering Cornell Univ. Ithaca, N.Y., 1969.

[5] P.A.P. Moran, The Theory of Storage, Methuen, London, 1959.

[6] A. Prékopa, Logarithmic Concave Measures with Application to Stochastic Programming, Acta Scientiarum Mathematicarum (Szeged) 35 (1971) 301-316.

[7] A. Prékopa, A Class of Stochastic Programming Decision Problems, Mathematische Operationsforschung und Statistik 3(1972) 349-354.

[8] A. Prékopa. Stochastic Programming Models for Inventiry Control and Water Storage Problems, Colloquia Mathematica Societatis János Bolyai, 7, Inventory Control and Water Storage. Győr (Hungary), 1971, North Holland Publishing Co. 1973, 229-245.

[9] A. Prékopa. Logarithmic Concave Measures and Related Topics, Proceedings of the International Conference on Stochastic Programming, Oxford, 1974, Acad. Press (to appear).

[10] A. Prékopa and T. Szántai, A New Multivariate Gamma Distribution and its Fitting to Empirical Data, Water Resources Research (to appear).

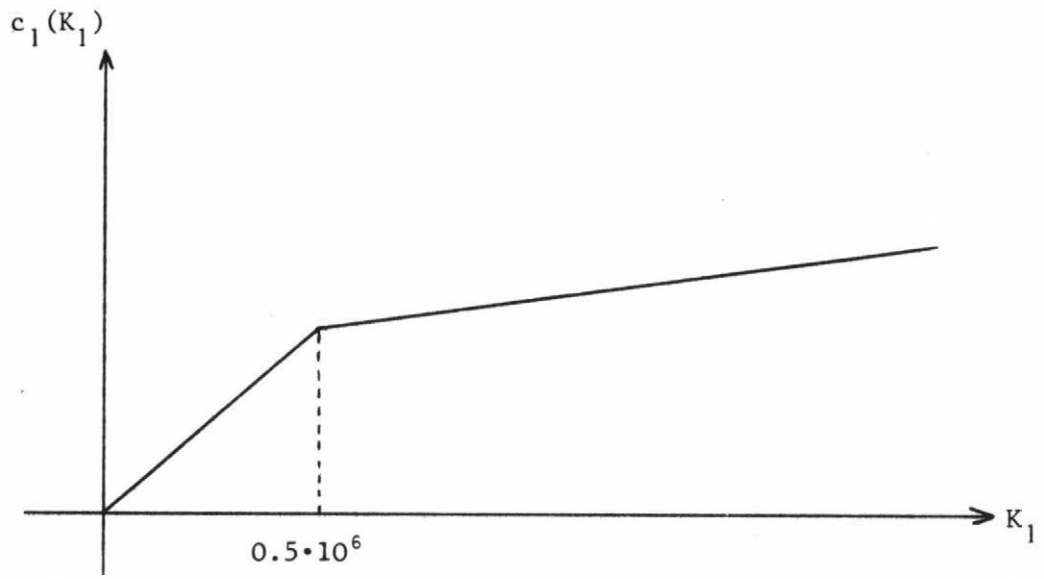


Figure 2.

Plot of the function $c_1(K_1)$. The function is concave on the halfline $K_1 \geq 0$ but it is linear and thus also convex on the halfline $K_1 \geq 0.5 \cdot 10^6$.

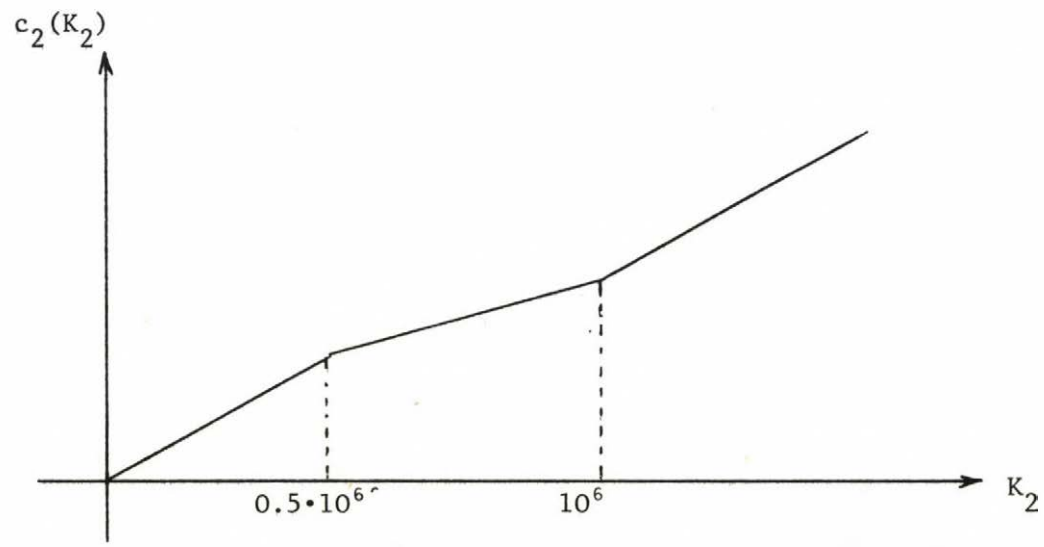


Figure 3.

Plot of the function $c_2(K_2)$. The function is convex on the halfline $K_2 \geq 0$.

A New Multivariate Gamma Distribution and its Fitting to Empirical Streamflow Data*

András Prékopa
Technical University of Budapest
and
Hungarian Academy of Sciences

Tamás Szántai
Technical University of Budapest

ABSTRACT

A new multivariate gamma distribution is presented which can sucessfully be fitted to empirical data where the one-dimensional marginal distributions are gamma-distributions with prescribed parameters and the correlations are nonnegative. It is not intended to give explicit formulae either for the joint density or for the joint characteristic function of the random variables. Our representation of the individual gamma-distributed random variables will be used for simulation with the aid of which we approximate probabilities of sets in higher dimensional spaces. Since streamflow and other hydrological data frequently follow gamma distribution and also they are frequently stochastically dependent our multivariate distribution and fitting technique seems to be of particular interest from the hydrological point of view.

1. INTRODUCTION

A multivariate probability distribution is said to be a multivariate gamma distribution if its one-dimensional marginal distributions are gamma distributions. Many such multivariate distributions were already defined in the literature. A bibliography is given at the end of the paper.

*This work was supported by the National Water Authority of Hungary.

A one-dimensional continuous probability distribution is said to be a gamma distribution if its probability density has the following form

$$(1.1) \qquad \frac{\lambda^{\nu} x^{\nu-1} e^{-\lambda x}}{\Gamma(\nu)} , \qquad x > 0,$$

where ν,λ are positive constants and $\Gamma(\nu)$ is the value of the well-known gamma function at ν, i.e.

$$(1.2) \qquad \Gamma(\nu) = \int_0^{\infty} x^{\nu-1} e^{-x} dx.$$

It is well-known that if a random variable ξ has the probability density (1.1), then

$$(1.3) \qquad E(\xi) = \frac{\nu}{\lambda} , \quad D^2(\xi) = \frac{\nu}{\lambda^2} ,$$

where E denotes expectation and D denctes standard deviation. The random variable $\zeta=\lambda\xi$ has to probability density

$$(1.4) \qquad \frac{x^{\nu-1} e^{-x}}{\Gamma(\nu)} , \qquad x > 0.$$

The probability distribution having density (1.4) is called standard gamma distribution with parameter ν. This probability distribution has the property that the expectation equals the variance as it turns out from the formula (1.4). The following property will be very important for us: if the random variables ζ_1 and ζ_2 have standard gamma distributions with parameter ν_1 and ν_2, respectively, then the random variable $\zeta_1+\zeta_2$ also has a standard gamma distribution and its parameter equals $\nu_1+\nu_2$.

Chereiyan [2] introduced first and investigated a two-dimensional gamma distribution which, according to his definition, is the joint distribution of the sums

$$\zeta_1 = \eta_o + \eta_1 \quad , \tag{1.5}$$
$$\zeta_2 = \eta_o + \eta_2 \quad ,$$

where η_o, η_1, η_2 are independent random variables having standard gamma distributions. Ramabhadran [16] generalized this to the n-dimensional case. His distribution is the joint distribution of the following sums

$$\zeta_i = \eta_o + \eta_i \quad , \quad i = 1,2,\ldots,n, \tag{1.6}$$

where η_o, η_1 ,..., η_n are independent and have standard gamma distributions.

We follow essentially this line but our objective in this paper is not so much the investigation of the mathematical properties of a multivariate gamma distribution; we give a fitting procedure and at the same time introduce a multivariate gamma distribution satisfying certain requirements.

The problem of defining and fitting a multivariate gamma distribution arises in the following manner. Suppose we have a finite number of empirical time series representing a sample for some random vector. Then we compute the empirical expectations and covariances. Suppose we also know that the individual random variables have separately gamma distributions. Now we wish to compute, either using an analytic method or some Monte Carlo technique, probabilities in higher dimensional spaces. To do this a multivariate probability distribution is needed which has the prescribed marginal distributions and convariances.

The gamma distribution is one of the favorite probability distributions with streamflow and other hydrological data. Mathematical theory supports its use in connection with streamflow data since in the theory of secondary stochastic processes it arises as the probability distribution superimposed exponential curves at some fixed time point with exponentially distributed amplitude

where the underlying random time point sequence forms a homogeneous Poisson process [17].

Nowadays it is fashionable the use of the lognormal distribution as the probability distribution of streamflow data. In the multivariate case it is then assumed that the logarithms of the random variables follow a multivariate normal distribution. This assumption can certainly be made if we only want to fit the distribution and disregard further considerations where the physical background is somehow involved. The lognormal distribution of fractions arising in griding or similar processes; the product of independent lognormally distributed random variables is also lognormally distributed whereas the sum is never. It is not an infinitely divisible distribution and the failure rate is not a monotonic function. Hence addition is physically unnatural in connection with lognormally distributed random variables and this practically excludes it from storage theory.

We hope that our multivariate gamma distribution provides a good alternative if we want to fit multivariate distribution to empirical data which are correlated and have skew marginal distributions.

2. DEFINITION OF THE NEW MULTI-GAMMA PROBABILITY DISTRIBUTION.

Let $\xi_1, \ldots, \xi_n$ be random variables having gamma distributions with parameters $\lambda_1, \Theta_1, \ldots, \lambda_n, \Theta_n$, respectively. Then the random variables $\lambda_1\xi_1, \ldots, \lambda_n\xi_n$ have standard gamma distributions with parameters $\Theta_1, \ldots, \Theta_n$, respectively. We wish to approximate their joint probability distribution by the joint distribution of the components $\zeta_1, \ldots, \zeta_n$ of a random vector

$$\underline{\zeta} = A\underline{\eta} \tag{2.1}$$

where $\underline{\eta}$ has independent and standard gamma distributed components and A is a matrix with 0,1 entries.

If we want to ensure that the expectations of $\lambda_1\xi_1, \ldots, \lambda_n\xi_n$ be the same as those of $\zeta_1, \ldots, \zeta_n$ and the covariances be as closely approximated

as is possible, then it is clear that it is good to choose the number of columns of the matrix A as large as is possible. The number of column vectors having 0 or 1 components equals the number of different combinations which can be chosen from n different elements. This number is 2^n. Disregarding of the column vector having all 0 components, it turns out that A should have 2^n-1 columns. If n = 4, then A is the following matrix

$$(2.2) \qquad A=\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

This means that the components of $\underline{\zeta}$ are represented in the following manner

$$\begin{aligned} \zeta_1 &= \eta_1 \quad +\eta_5+\eta_6+\eta_7 \quad +\eta_{11}+\eta_{12}+\eta_{13} \quad +\eta_{15}, \\ \zeta_2 &= \eta_2 \quad +\eta_5 \quad +\eta_8+\eta_9 \quad +\eta_{11}+\eta_{12} \quad +\eta_{14}+\eta_{15}, \\ \zeta_3 &= \eta_3 \quad +\eta_6 \quad +\eta_8 \quad +\eta_{10}+\eta_{11} \quad +\eta_{13} \quad +\eta_{14}+\eta_{15}, \\ \zeta_4 &= \eta_4 \quad +\eta_7 \quad +\eta_9+\eta_{10} \quad +\eta_{12}+\eta_{13} \quad +\eta_{14}+\eta_{15}. \end{aligned} \qquad (2.3)$$

We remark that the convarianve of partial sums of independent random variables is the sum of variances of the common terms. Taking this into account it follows that

a.) $\lambda_1\xi_1, \ldots, \lambda_n\xi_n$ have the same one-dimensional distributions, as the random variables $\zeta_1, \ldots, \zeta_n$, respectively,

b.) the covariances of the two random vectors having the above components coincide if and only if the following conditions are satisfied

$$(2.4) \qquad A\underline{\nu} = \underline{\Theta} \ ,$$

$$(2.5) \qquad \tilde{A}\underline{\nu} = \underline{c} \ ,$$

$$(2.6) \qquad \underline{\nu} \geq \underline{0} \ ,$$

where $\underline{\Theta}$ is the vector of components $\Theta_1,\ldots,\Theta_n$ $\underline{\nu}$ is the vector of the unknown components $\nu_1,\ldots,\nu_r$ $(r=2^n-1)$ which are the parameters of the random variables $\eta_1,\ldots,\eta_r$, $\underline{c}$ is a vector containing all covariances of $\lambda_1\xi_1,\ldots,\lambda_n\xi_n$ in the ordering given below:

(2.7)
$$\underline{c}' = (c_{11},\ldots,c_{nn},c_{12},\ldots,c_{1n},c_{23},\ldots,c_{2n},\ldots,c_{n-1,n})$$

and $\tilde{A}$ is a matrix of order $[n(n+1)/2] \times r$ constructed of the componentwise products of the rows of A and these rows follow each other in accordance with the ordering of the components in the vector $\underline{c}$.

We observe the Equation (2.4) is superfluous because it is contained among the equations (2.5) if we select those where on the right hand side there stand $c_{11},\ldots,c_{nn}$. In fact $c_{11}=\Theta_1,\ldots,c_{nn}=\Theta_n$ and the elementwise products of the rows of A with themselves coincide with the original rows.

The equation contained in (2.5) are detailed below for the case n=4:

(2.8)
$$\begin{aligned}
&\nu_1 + \nu_5 + \nu_6 + \nu_7 + \nu_{11} + \nu_{12} + \nu_{13} + \nu_{15} = c_{11},\\
&\nu_2 + \nu_5 + \nu_8 + \nu_9 + \nu_{11} + \nu_{12} + \nu_{14} + \nu_{15} = c_{22},\\
&\nu_3 + \nu_6 + \nu_8 + \nu_{10} + \nu_{11} + \nu_{13} + \nu_{14} + \nu_{15} = c_{33},\\
&\nu_4 + \nu_7 + \nu_9 + \nu_{10} + \nu_{12} + \nu_{13} + \nu_{14} + \nu_{15} = c_{44},\\
&\nu_5 + \nu_{11} + \nu_{12} + \nu_{15} = c_{12},\\
&\nu_6 + \nu_{11} + \nu_{13} + \nu_{15} = c_{13},\\
&\nu_7 + \nu_{12} + \nu_{13} + \nu_{15} = c_{14},\\
&\nu_8 + \nu_{11} + \nu_{14} + \nu_{15} = c_{23},\\
&\nu_9 + \nu_{12} + \nu_{14} + \nu_{15} = c_{24},\\
&\nu_{10} + \nu_{13} + \nu_{14} + \nu_{15} = c_{34}.
\end{aligned}$$

To the representation (2.1) we make the following remarks.

Remark 1. Though in the representation $\underline{\zeta} = A\underline{\eta}$ we use 2^n-1 random variables on the right hand side, it is always true that if there exists a $\underline{\nu}$ satisfying (2.5) and (2.6), then there exists at least one $\underline{\nu}$ such that it satisfies the same relations and the number of components different from zero is at most $n(n+1)/2$. This follows immediately from the well-known theorem in linear programming stating that if a linear equality system has a nonnegative (feasible) solution then it also has a basic nonnegative (feasible) solution.

Note that $\tilde{A}$ always contains a unit matrix. This is due to the fact that A contains n columns which are unit vectors and $\binom{n}{2}$ different columns consisting of n-2 zeros and 2 ones. Thus the rank of $\tilde{A}$ equals $n(n+1)/2$ and this is at the same time the number of basic components in case of any basis.

It follows that we do not need more then $n(n+1)/2$ random variables on the right hand side of (2.1). This is very important because $n(n+1)/2$ increases much more slowly than 2^n-1 and if we want to use the representation (2.1) for simulation then we may keep the number of operations below a realistic level.

Remark 2. Conditions (2.5) and (2.6) may uniquely determine $\underline{\nu}$. If this is not the case, i.e. there exist at least two different such $\underline{\nu}$ vectors which satisfy (2.5) and (2.6), then there exist infinitely many $\underline{\nu}$'s which satisfy (2.5) and (2.6). Thus in general the multivariate gamma distribution satisfying our requirements is not uniquely determined.

Remark 3. It is not always possible to satisfy Relations (2.5) and (2.6). In fact if n=2, then these relations are the following:

$$
\begin{aligned}
&\nu_1 + \nu_2 = c_{11}, \\
&\nu_1 + \nu_3 = c_{22}, \\
&\nu_1 = c_{12}, \\
&\nu_1 \geq 0, \ \nu_2 \geq 0, \ \nu_3 \geq 0,
\end{aligned}
\tag{2.9}
$$

and if we have the following covariance matrix

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$$

then (2.9) clearly cannot be satisfied. In the next section we shall give methods with the aid of which we at least approximate the joint distribution of $\lambda_1\xi_1, \ldots, \lambda_n\xi_n$ by the joint distirbution of $\zeta_1, \ldots, \zeta_n$. If n is large enough, however, then we hope that a $\underline{\nu}$ satisfying (2.5) and (2.6) exists because $\tilde{A}$ has a large number of columns as compared to the number of rows.

Remark 4. In a similar way we can define other multivariate probability distributions. We only have to require that $E(\eta_i) = D^2(\eta_i)$. E.g. we can define and fit to empirical data multivariate Poission distribution.

3. FITTING OF THE MULTIGAMMA DISTRIBUTION TO EMPIRICAL DATA

We want to find at least one $\underline{\nu}$ satisfying Relations (2.5) and (2.6). This can numerically be done e.g. with the aid of the first phase of linear programming. At the end of the first phase either we have a $\underline{\nu}$ satisfying the requirements or it turns out that such a $\underline{\nu}$ does not exist. If this is the case, then we want to find a $\underline{\nu}$ such that $\underline{\nu} \geq \underline{0}$ and it nearly satisfies (2.5). We shall formulate three different measures of deviation. These measures are constructed so that every one of them is equal to zero if and only if (2.5) and (2.6) can be satisfied. Below we describe them measures and their minimization techniques which are the fitting procedures of the multigamma distribution to empirical data.

Minimizing the sum of absolute deviations between the two sides in (2.5).

We introduce the $n(n+1)/2$ -dimensional vectors $\underline{u},\underline{v}$ and formulate the following problem:

$$\text{minimize} \quad \left(\sum_{i=1}^{\frac{1}{2}n(n+1)} u_i + \sum_{i=1}^{\frac{1}{2}n(n+1)} v_i \right)$$

(3.1) subject to the constraints

$$\underline{u} - \underline{v} + \tilde{A}\,\underline{\nu} = \underline{c} \ ,$$

$$\underline{u} \geq \underline{0} \ , \quad \underline{v} \geq \underline{0} \ , \quad \underline{\nu} \geq \underline{0}.$$

The optimum value of the objective function is zero if and only if (2.5) and (2.6) can be satisfied. Since $\underline{c}$ has all nonnegative components, it follows that $\underline{u} = \underline{c}$, $\underline{v} = \underline{0}$, $\underline{\nu} = \underline{0}$ can be used as an initial solution to the linear programming problem (3.1).

It may turn out that for Problem (3.1) such an optimal solution $\underline{u}^*$, $\underline{v}^*$, $\underline{\nu}^*$ is obtained from which $\underline{\nu}^*$ does not satisfy (2.4). We recall that (2.4) is contained among the constraints (2.5) and the validity of (2.4) is necessary and sufficent that our fitted multivariate distribution has the prescribed one-dimensional marginal distributions. If we insist that this last property must hold then (3.1) has to be modified in the following manner: we fix $u_1=\ldots=u_n=v_1=\ldots=v_n=0$ and allow to be variable only the remaining components of $\underline{u}$ and $\underline{v}$. If $n = 3$ then our new problem reads as follows:

$$\text{minimize} \quad (u_4+u_5+u_6+v_4+v_5+v_6)$$

subject to the constraints

$$
\begin{array}{cccccccccccc}
 & & & & & & \nu_1 & & +\nu_4 & +\nu_5 & & +\nu_7 & = c_{11} \\
 & & & & & & & \nu_2 & +\nu_4 & & +\nu_6 & +\nu_7 = c_{22} \\
 & & & & & & & & \nu_3 & & +\nu_5 & +\nu_6 & +\nu_7 = c_{33} \\
u_4 & & -v_4 & & & & & & \nu_4 & & & +\nu_7 & = c_{12} \\
 & u_5 & & -v_5 & & & & & & \nu_5 & & +\nu_7 & = c_{13} \\
 & & u_6 & & -v_6 & & & & & & \nu_6 & +\nu_7 & = c_{14}
\end{array}
\tag{3.2}
$$

$u_i \geq 0$, $v_i \geq 0$, $i=1,2,3$; $\nu_i \geq 0$, $i=1,2,3,4,5,6,7$.

As an initial basic feasibles solution we can choose the following $\nu_1 = c_{11}$, $\nu_2 = c_{22}$, $\nu_3 = c_{33}$, $u_4 = c_{12}$, $u_5 = c_{13}$, $u_6 = c_{14}$. In general Problems of the type (3.2) have such a matrix on the left hand side in the equality constraints which contain a unit matrix.

Minimizing the sum of squares of the deviations between the two sides in (2.5).

Now the fitting procedure consists in the solution of the following quadratic programming problem:

$$\text{minimize} \quad \sum_{i=1}^{\frac{1}{2}n(n+1)} u_i^2 \quad ,$$

subject to the constraints

$$
\begin{aligned}
\underline{u} + \tilde{A}\,\underline{\nu} &= \underline{c} \ , \\
\underline{\nu} &\geq \underline{0} \ .
\end{aligned}
\tag{3.3}
$$

According to the Kuhn-Tucker theorem [12], $\underline{u}$ and $\underline{v} \geq \underline{0}$ minimize the objective function of Problem (3.3) if and only if there exist $\underline{\lambda}$ and $\underline{v} \geq \underline{0}$ vectors such that

$$
\begin{aligned}
\underline{u} + \tilde{A}\underline{v} &= \underline{c}\,, \\
-2\underline{u} - \underline{\lambda} &= \underline{0}\,, \\
-\tilde{A}'\underline{\lambda} + \underline{v} &= \underline{0}\,, \\
\underline{v}'\underline{v} &= 0.
\end{aligned}
\tag{3.4}
$$

The system of equalities (3.4) has the following equivalent form

$$
\begin{aligned}
\underline{u} + \tilde{A}\underline{v} &= \underline{c}\,, \\
2\tilde{A}'\underline{u} + \underline{v} &= \underline{0}\,, \\
\underline{v}'\underline{v} &= 0.
\end{aligned}
\tag{3.5}
$$

Here in the first two rows the equality system has a matrix of order

$$\left[\frac{1}{2}\, n(n+1) + 2^n - 1\right] \times \left[\frac{1}{2}\, n(n+1) + 2(2^n - 1)\right].$$

Both the row and column numbers are very large even in case of a relatively moderate n hence this fitting method is not practical enough.

Minimizing the maximal absolute deviation between the two sides in (2.5).

In this case we want to minimize y subject to the constraints

$$\begin{aligned} \tilde{A}\underline{\nu} - \underline{c} &\leq \underline{1}y \,, \\ -\tilde{A}\underline{\nu} + \underline{c} &\leq \underline{1}y \,, \end{aligned} \tag{3.6}$$

where $\underline{1}$ is the vector all components of which are equal to 1. This problem is equivalent to the following linear programming problem

$$\text{minimize } y$$

subject to the contraints

$$\begin{aligned} \underline{u} \quad & +\tilde{A}\underline{\nu} - \underline{1}y = \underline{c} \,, \\ \underline{v} \quad & -\tilde{A}\underline{\nu} - \underline{1}y = -\underline{c} \,, \\ & \underline{u} \geq \underline{0} \,, \ \underline{v} \geq \underline{0} \,, \ \underline{\nu} \geq \underline{0} \,, \ y \geq 0. \end{aligned} \tag{3.7}$$

This problem is relatively simple because the number of equality constraints is not much larger then that of (2.5) ; it only doubles.

An initial feasible basis can be obtained in the following manner. We choose the columns of the variables contained in $\underline{u}$ and $\underline{v}$. Then find the most negative component of $-\underline{c}$ and if this is $-c_i$, then we replace the column of v_i by the column of the variable y, containing all -1 components. To this basis the corresponding basis vector can be found in a trivial way.

It is practical to apply the revised simplex method for the solution of Problems (3.1) and (3.7) because the matrix of the equality constraints is sparse further it is not necessary in this case to store columns of $\tilde{A}$. We only generate them in case of necessity. Since this fitting procedure can be carried out very effectively it seems to be realistic to say that using these methods we can also fit our multigamma

distribution in the case when the number of components is large, say $n = 100$ or n is even somewhat larger.

4. CONDITIONAL DISTRIBUTIONS.

Let us consider the random vector $\underline{\zeta} = A\underline{\eta}$. In this section we duscuss the joint probability distribution of the components of $\underline{\zeta}$ given that the value of one of them say, ζ_1, is fixed.

First of all we note the very important property that the components of $\underline{\zeta}$ are stochastically independent if and only if they are uncorrelated. This follows immediately from the fact that the covariance of ζ_i is the sum of variances of the joint terms in the representation $\underline{\zeta} = A\,\underline{\eta}$.

Another important fact, that we shall use in the sequel, is that if $\eta_1, \ldots, \eta_q$ are independent random variables having standard gamma distributions, then the random vector with components

$$\frac{\eta_1}{\eta_1 + \ldots + \eta_q}, \ldots, \frac{\eta_q}{\eta_1 + \ldots + \eta_q}$$

is independent of the random variable

$$\eta_1 + \ldots + \eta_q .$$

For the proof of this fact see e.g. [19].

Let $r = n(n+1)/2$ as before and denote

$$\zeta_2^{(1)}, \ldots, \zeta_r^{(1)},$$

respectively, the sum of joint η terms of $\zeta_2, \ldots, \zeta_r$ and ζ_1. Let further

$$\zeta_2^{(2)} = \zeta_2 - \zeta_2^{(1)} \quad , \ldots, \quad \zeta_r^{(2)} = \zeta_r - \zeta_r^{(1)} \ .$$

The random variables

$$(4.1) \qquad \frac{\zeta_2^{(1)}}{\zeta_1} \ , \ \ldots \ , \ \frac{\zeta_r^{(1)}}{\zeta_1}$$

have separately beta distributions, the probability density of $\zeta_i^{(1)} / \zeta_1$ is (see [19])

$$\frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} x^{\alpha_i - 1} (1-x)^{\beta_i - 1} \ , \quad 0 < x < 1, \quad i = 2, \ldots, r,$$

where

$$\alpha_i = E(\zeta_i^{(1)}) \ , \quad \beta_i = E(\zeta_i^{(2)}) \ , \quad i = 2, \ldots, r,$$

and the random vector of components (4.1) is independent of ζ_1. In wiev of these we can write

$$\begin{aligned} & P(\zeta_2 < z_2 \ , \ldots, \ \zeta_r < z_r \mid \zeta_1 = z_1) = \\ & = P(\zeta_2^{(1)} + \zeta_2^{(2)} < z_2 \ , \ldots, \ \zeta_r^{(1)} + \zeta_r^{(2)} < z_r \mid \zeta_1 = z_1) = \\ & = P(z_1 \frac{\zeta_2^{(1)}}{\zeta_1} + \zeta_2^{(2)} < z_2 \ , \ldots, \ z_1 \frac{\zeta_r^{(1)}}{\zeta_1} + \zeta_r^{(2)} < z_r \mid \zeta_1 = z_1) = \\ & = P\left(z_1 \frac{\zeta_2^{(1)}}{\zeta_1} + \zeta_2^{(2)} < z_2 \ , \ldots, \ z_1 \frac{\zeta_r^{(1)}}{\zeta_1} + \zeta_r^{(2)} < z_r\right). \end{aligned} \qquad (4.2)$$

Thus the conditional distribution equals the distribution of a sum $z_1 \underline{\beta} + \underline{\gamma}$ where $\underline{\beta}$ is the r-1-dimensional random vector having the components (4.1) and $\underline{\gamma}$ is the random vector of components $\zeta_2^{(2)}, \ldots, \zeta_r^{(2)}$. $\underline{\gamma}$ has the multigamma distribution introduced in this paper and the components of $\underline{\beta}$ are similarly constructed but instead of independent standard gamma variables we now have partial sums of components of a random vector having Dirichlet distribution [19].

The expectation of $z_1 \underline{\beta} + \underline{\gamma}$ is a linear function of z_1 hence the regression of ζ_i $(i \neq 1)$ with respect to ζ_1 is lienar.

5. NUMERICAL EXAMPLES.

For the monthly streamflows of the Tisza river at Tokaj a 60 year time series exists. We consider six months: April, May, June, July, August, September and want to fit a multigamma distribution to the empirical data. We have the following empirical expectations, standard deviations (in 10^6 m^3) and correlations:

	Expectation	Standard deviation
April	2337.21	1110.72
May	1725.96	958.27
June	1095.64	522.99
July	985.96	969.68
August	734.24	553.40
September	728.44	768.59

$$R = \begin{pmatrix} 1.000 & 0.646 & 0.317 & 0.000 & 0.070 & 0.174 \\ & 1.000 & 0.532 & 0.229 & 0.201 & 0.139 \\ & & 1.000 & 0.437 & 0.284 & 0.334 \\ & & & 1.000 & 0.746 & 0.274 \\ & & & & 1.000 & 0.382 \\ & & & & & 1.000 \end{pmatrix}$$

Let ξ_1 , ξ_2 , ξ_3 , ξ_4 , ξ_5 , ξ_6 denote the above random monthly streamflows. From the empirical data it turns out that each of them can be assumed to have gamma distribution. We obtained the following parameter values:

$\lambda_1 = 0.0018945,$ $\Theta_1 = 4.4277$,

$\lambda_2 = 0.0018796,$ $\Theta_2 = 3.2441$,

$\lambda_3 = 0.0040057,$ $\Theta_3 = 4.3888$,

$\lambda_4 = 0.0020314,$ $\Theta_4 = 2.0029$,

$\lambda_5 = 0.0023975,$ $\Theta_5 = 1.7603$,

$\lambda_6 = 0.0012331,$ $\Theta_6 = 0.8983$.

As a result of a fitting procedure where we minimized the sum of absolute deviations, we found that Relations (2.5) and (2.6) can be satisfied. Many ν_i are equal to zero in the optimal solution. In order to avoid writing very long expressions for the ζ_i random variables, we gave a new numbering for the η_i and ν_i. The parameter values and the representations given below use this new numbering. The results are:

$\nu_1 = 0.06794,$ $\nu_2 = 0.70346,$ $\nu_3 = 1.48655,$ $\nu_4 = 0.20576,$

$\nu_5 = 0.21985,$ $\nu_6 = 0.28987,$ $\nu_7 = 0.00285,$ $\nu_8 = 0.00309,$

$\nu_9 = 0.11489,$ $\nu_{10} = 0.09458,$ $\nu_{11} = 0.39688,$ $\nu_{12} = 0.10628,$

$\nu_{13} = 0.07771,$ $\nu_{14} = 1.67685,$ $\nu_{15} = 0.12304,$ $\nu_{16} = 0.17998,$

$\nu_{17} = 1.03670,$ $\nu_{18} = 0.14979,$ $\nu_{19} = 0.04488,$ $\nu_{20} = 1.21645,$

$$\zeta_1 = \eta_{14}+\eta_{15}+\eta_{16}+\eta_{17}+\eta_{18}+\eta_{19}+\eta_{20}$$

$$\zeta_2 = \eta_7+\eta_8+\eta_9+\eta_{10}+\eta_{11}+\eta_{12}+\eta_{13} \quad +\eta_{17}+\eta_{18}+\eta_{19}+\eta_{20}$$

$$\zeta_3 = \eta_3+\eta_4+\eta_5+\eta_6 \quad +\eta_9+\eta_{10}+\eta_{11}+\eta_{12}+\eta_{13} \quad +\eta_{16} \quad +\eta_{20}$$

$$\zeta_4 = \eta_2 \quad +\eta_4+\eta_5+\eta_6 \quad +\eta_8 \quad +\eta_{11}+\eta_{12}+\eta_{13}$$

$$\zeta_5 = \eta_1+\eta_2 \quad +\eta_5+\eta_6+\eta_7+\eta_8 \quad +\eta_{10} \quad +\eta_{12}+\eta_{13} \quad +\eta_{18}+\eta_{19}$$

$$\zeta_6 = \eta_1 \quad +\eta_6 \quad +\eta_9 \quad +\eta_{13} \quad +\eta_{15}+\eta_{16} \quad +\eta_{19}$$

The joint probability distribution of $\zeta_2,\zeta_3,\zeta_4,\zeta_5,\zeta_6$ given that $\zeta_1=z_1$ is the same as the joint probability distribution of the following random variables:

$$\delta_2 = \eta_7+\eta_8+\eta_9+\eta_{10}+\eta_{11}+\eta_{12}+\eta_{13}+ z_1 \frac{\eta_{17}+\eta_{18}+\eta_{19}+\eta_{20}}{\eta_{14}+\eta_{15}+\eta_{16}+\eta_{17}+\eta_{18}+\eta_{19}+\eta_{20}},$$

$$\delta_3 = \eta_3+\eta_4+\eta_5+\eta_6+\eta_9+\eta_{10}+\eta_{11}+\eta_{12}+\eta_{13}+z_1 \frac{\zeta_{16}+\zeta_{20}}{\eta_{14}+\eta_{15}+\eta_{16}+\eta_{17}+\eta_{18}+\eta_{19}+\eta_{20}},$$

$$\delta_4 = \eta_2+\eta_4+\eta_5+\eta_6+\eta_8+\eta_{11}+\eta_{12}+\eta_{13},$$

$$\delta_5 = \eta_1+\eta_2+\eta_5+\eta_6+\eta_7+\eta_8+\eta_{10}+\eta_{12}+\eta_{13}+z_1 \frac{\eta_{18}+\eta_{19}}{\eta_{14}+\eta_{15}+\eta_{16}+\eta_{17}+\eta_{18}+\eta_{19}+\eta_{20}},$$

$$\delta_6 = \eta_1+\eta_6+\eta_9+\eta_{13}+ z_1 \frac{\eta_{15}+\eta_{16}+\eta_{19}}{\eta_{14}+\eta_{15}+\eta_{16}+\eta_{17}+\eta_{18}+\eta_{19}+\eta_{20}}.$$

These formulae enable effective simulation for the conditional joint probability distribution of our random variables. In fact this simulation goes back to the simulation of independent gamma variables for which effective methods are known in the literature, see e.g. [1].

REFERENCES

[1] J.H. Ahrens and V. Dieter, "Computer methods for sampling from gamma, beta, Poisson and binominal distributions", Computing 12(1974) 223-246.

[2] K.C. Cheriyan, "A bivariate correlated gamma-type distribution function", Journal of the Indian Mathematical Society 5(1941) 133-144.

[3] G.B. Dantzig, "Linear Programming and Extensions (Princeton University Press, Princeton, New Yersey, 1963).

[4] D.P. Gaver, "Multivariate gamma distributions generated by mixture" Sankhya, Series A 32(1970) 123-126.

[5] D.R. Jensen, "The joint distribution of traces of Wishart matrices and some applications" Annales of Mathematical Statistics 41(1970) 133-145.

[6] N.L. Johnson, "Some notes on the investigation of heterogeneity in interactions", Trabajos de Estadistica 13(1962). 183-199.

[7] N.L. Johnson and S. Kotz, "Distributions in Statistics. Continuous Multivariate Distributions (John Wiley, New York, 1972).

[8] W.F. Kibble, "A two-variate gamma type distribution " Sankhya, Series A 5(1941) 137-150.

[9] P.R. Krishnaiah, P. Hagis and L. Steinberg, "A note on the bivariate chi-distribution", SIAM Review 5(1963) 140-144.

[10] P.R. Krishnaiah and M.M. Rao, "Remarks on a multivariate gamma distribution", American Mathematical Monthly 68(1961) 342-346.

[11] A.S. Krishnamorthy and M. Parthasarathy, "A multivariate gamma-type distribution", Annales of Mathematical Stat. 22(1951) 549-557 (corrections 31(1969) 229.).

[12] H. W. Kuhn and A. W. Tucker, "Nonlinear programming" in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (University of California Press, Berkeley, California, 1950) 481-492.

[13] E. Lukacs and L.G. Laha, Applications of Characteristic Functions (Griffin's Statistical Monograph No. 14, Griffin, London, 1964).

[14] P. A. P. Moran, "Testing for correlation between non-negative variates", Biometrika 54(1967) 385-394.

[15] P. A. P. Moran and D. Vere-Jones, "The infinite divisibility of multivariate gamma distributions", Sankhya, Series A 31(1969) 191-194.

[16] V. R. Ramabhadran, "A multivariate gamma-type distribution", Sankhya, Series A 11(1951) 45-46.

[17] L. Takács, "On secondary processes generated by a Poisson Process and their applications in physics", Acta Math. Acad. Sci. Hung. 5(1954) 203-236.

[18] D. Vere-Jones, "The infinite divisibility of a bivariate gamma distirbution", Sankhya, Series A 27(1967) 421-422.

[19] S.S. Wilks, Mathematical Statistics , Wiley, Now York, London, 1962.

[20] Ph. Wolfe, "The simplex method for quadratic programming", Econometica 27(1959).

[21] V.K. Zaharov, O.V. Sarmanov and B.A. Sevast'janov, "A sequential 2 test", Mat. Sb.(N.S.) 79 (121) (1969), 444-460, (in Russian).

[22] I.O.Sarmanov, "Generalized symmetric gamma-correlation", Dokl. Akad. Nauk SSSR 179 (1968) 1279-1281, (in Russian).

[23] I.O. Sarmanov, "A gamma-correlation process and its properties", Dokl. Akad. Nauk SSSR 191(1970) 30-32, (in Russian).

ON OPTIMAL REGULATION OF A STORAGE LEVEL WITH APPLICATION TO TO THE WATER LEVEL REGULATION OF A LAKE

András Prékopa

and

Tamás Szántai

ABSTRACT

In the paper [10] a system of stochastic programming was introduced for the optimal control of a storage level. Each model in this system serves to determine the optimal for only one period ahead though the time horizon consists of many future periods. The optimal control thus obtained can be considered an open loop control methodology. The main purpose of this paper is to present an application by given an optimal control method for the regulation of the water level of Lake Balaton in Hungary. By solving almost 600 stochastic programming problems we analyze what would have happened if we had controlled the water level using our method between 1922 and 1970, where one decision period is one month. The numerical results show that the proposed control methodology works quite well in this case.

1. INTRODUCTION.

In the paper [10] a system of stochastic programming models was introduced for the optimal control of a storage level. Each model in this system serves to determine the optimal policy for only one period ahead though the time horizon consists of many future periods.
The optimal control thus obtained can be considered an open loop control by giving an optimal control method for the regulation of the water level of Lake Balaton in Hungary. By solving almost 600 stochastic programming problems we analyze what would have happened if we hand controlled the water

level using our method between 1922 and 1970, where one decision period is one month. The numerical results show that the proposed control methodology works quite well in this case.

Our water input stochastic process will be assumed to be Gaussian. No time homogenity or independence, Markovian character or whatsoever will be supposed. Also the Gaussian nature of the input process is not an essential feature of our control methodology. We refer to the paper [11] where we used a multivariate gamma distribution introduced by the authors of this paper in connection with certain reservoir system operation model. It is possible to use also here the same multigamma or some other multivariate probabaility distribution. The case of the multivariate normal distribution - the application of which is supported by statistics in connection with Lake Balaton - is relatively simple because of the special properties of this multivariate probability density.

2. SHORT DESCRIPTION OF THE DYNAMIC MODEL SYSTEM USED FOR THE CONTROL OF THE STORAGE LEVEL.

Taking into account that application of our control methodology which we are going to present in the further sections of this paper, we shall use the terms corresponding to the water level regulation of a lake.

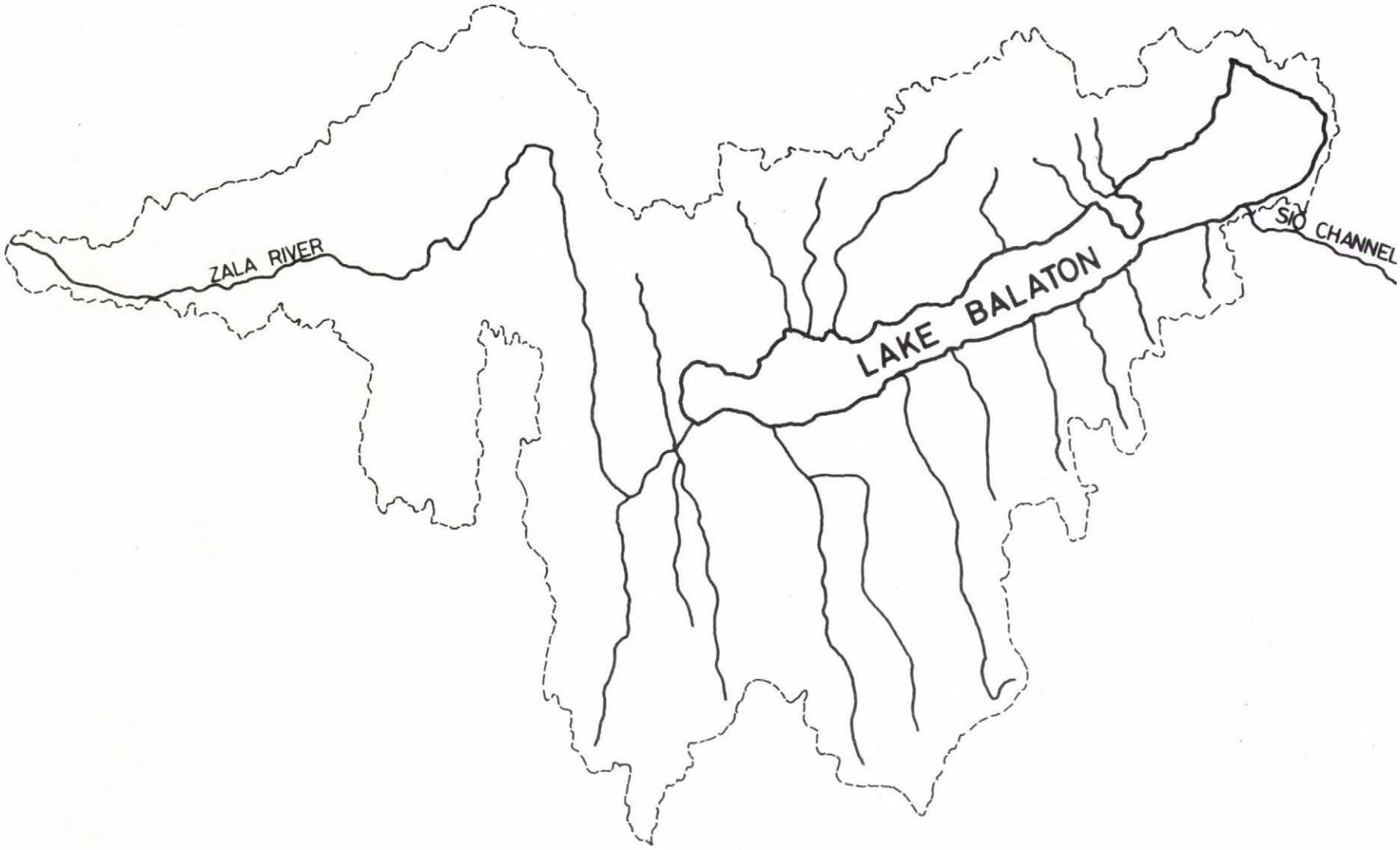


Fig. 1.
Lake Balaton and catchment area in the western part of Hungary
Water is released through Sio channel into Danube river

Figure 1. illustrates Lake Balaton in the western part of Hungary. Small rivers and rainfall represent positive inputs whereas evaporation represents negative input. The sum of these with positive resp. negative signs will be considered the water input. Thus the variables which denote water inputs in the subsequent periods can take on negative values too. Sometimes this in fact occurs.

Water can be released through the Sio channel into the Danube river. The monthly water quantities which can be released are limited by the capacity of the Sio channel. This capacity will be denoted symbolically by K in this section.

Instead of water levels we shall speak about water quantities. The connection relative to Lake Balaton between these two notions will be clarified later on. Let ζ_o be the initial water content of the lake and $\xi_1, \xi_2, \ldots$ the monthly random water inputs. ζ_o will be assumed to be nonrandon in our models. Let further $z_1, z_2, \ldots$ be decision variables belonging to the subsequent periods. These are the water quantities to be released through the channel in the subsequent periods. We decide on z_1 in the beginning of the first period, on z_2 in the beginning of the second period, etc. Introduce the notations

$$\zeta_k = \zeta_o + \xi_1 + \ldots + \xi_k, \qquad k = 1,2,\ldots$$

$$Z_k = z_1 + \ldots + z_k ,$$

The random process $\xi_1, \xi_2, \ldots$ will be assumed to be Gaussian. We prescribe further lower bounds $a_1, a_2, \ldots$ resp. upper bounds $b_1, b_2, \ldots$ for the water quantities being in the lake at the end of the subsequent periods. We consider the situation favourable if the inequalities

$$(2.1) \qquad a_k \leq \zeta_k - Z_k \leq b_k, \qquad k=1,2,\ldots$$

are satisfied, where the water quantities to be released $z_1, z_2, \ldots$ are subject to the inequalities

(2.2) $$0 \leq z_k \leq K, \qquad k = 1,2,\ldots$$

Since $\xi_1,\xi_2,\cdots$ are random variables, the fulfilment of the inequalities (2.1) cannot be guaranteed with probability 1. Before formulating our decision principle we make a remark concerning stochastic programming model construction.

Stochastic programming problems are formulated in such a way that first we formulate a deterministic mathematical programming problem, which is called underlying deterministic problem, then observe that some of the parameters in this problem are random in reality; in view of this problem looses its original meaning, hence we formulate another decision principle by taking into account the probability distribution of the random variables involved. Underlying mathematical programming problems can be either minimization (resp. maximization) problems or problems where we only wish to find at least one vector satisfying certain constraints. In this latter case the advised stochastic programming decision principle is to find that vector which maximizes the probability of the fulfilment of the random constraints subject to those constraints which do not contain random variables.

Inequalities (2.1) and (2.2) represent an underlying deterministic problem where we want to find $z_1,z_2,\ldots$ such that the referred inequalities be satisfied. The number of the considered periods should be finite. Having observed that $\xi_1,\xi_2,\ldots$ are random variables, we formulate the stochastic programming decision principle, in accordance with the above remark, so that we maximize the probability of the fulfilment of the inequalities (2.1) subject to the constraints (2.2).

Since we have the possibility to use a dynamic type decision methodology i.e. we have the possibility to decide in every period, the above mentioned principle will be turned into a sequence of problems and conditional probabilities will be maximized where in the condition there stand the already realised values of the stochastic process $\xi_1,\xi_2,\ldots$. The first problem in this sequence is the following

$$\text{maximize} \qquad P(a_k \leq \zeta_k - Z_k \leq b_k, \quad k = 1,\ldots,N)$$

(2.3) subject to

$$0 \leq z_k \leq K, \qquad k = 1,\ldots,N.$$

Out of the optimal solution $z_1^*, \ldots, z_N^*$ we only accept z_1^* and formulate the next problem. For the sake of simplicity the asterisk will be ommitted except for Section 5. Assume that we already fixed $z_1,\ldots,z_n$. Then in order to fix z_{n+1}, we formulate the following nonlinear programming problem

(2.4) $$\text{maximize} \quad P(a_k \leq \zeta_k - Z_k \leq b_k, \quad k=n+1,\ldots,n+N \mid \xi_1,\ldots,\xi_n)$$

subject to

$$0 \leq z_k \leq K, \quad k = n+1,\ldots,n+N.$$

Here $z_{n+1}, \ldots, z_{n+N}$ are the decision variables. Having computed the optimal solution, we only accept z_{n+1} as a final value. Thus our control methodology is fixed.

It should be mentioned that a positive lower bound for K may be required. Mathematically this does not present any difficulty. In fact if K_o is a positive lower bound for the z_k then using the new variables $y_k = z_k - K_o$, we can transform our problem into the already introduced form (2.3), (2.4).

3. MATHEMATICAL PROPERTIES OF THE MODEL SYSTEM INTRODUCED IN SECTION 2.

Before starting to discuss the subject mentioned in the title of this section, we recall some facts concerning logarithmic concave measures.

A nonnegative function $f(\underline{x})$, $\underline{x} \in R^m$ is said to be a logarithmic concave (point) function if for every $\underline{x}_1$, $\underline{x}_2 \in R^m$ and $0 < \lambda < 1$, we have

$$f(\lambda \underline{x}_1 + (1-\lambda)\underline{x}_2) \geq [f(\underline{x}_1)]^{\lambda} [f(\underline{x}_2)]^{1-\lambda}.$$

A measure P defined on the measurable subset of the space R^m is said to be logarithmic concave if for every pair A,B of convex subsets of R^m and $0<\lambda<1$, we have

$$P(\lambda A + (1-\lambda)B) \geq [P(A)]^{\lambda} [P(B)]^{1-\lambda}.$$

Here the sign $+$ denotes Minkowski addition i.e. in connection with two sets D,G, $D+G=\{\underline{d}+\underline{g} \mid \underline{d} \in D, \ \underline{g} \in G\}$, further the constant multiple λG of the set G is defined by the equality $G=\{\lambda \underline{g} \mid \underline{g} \in G\}$.

In [8] , [9] the following theorem was proved.

Theorem 1. If a probability measure P is generated by a logarithmic concave probability density i.e. for every measurable set $C \subset R^m$ we have

$$P(C) = \int_C f(\underline{x}) \, d\underline{x} ,$$

then P is a logarithmic concave measure.

Theorem 1 implies that if A is a convex subset of R^m then

$$\int_{A+\underline{x}} f(\underline{t}) \, d\underline{t}$$

is a logarithmic concave function of the variable $\underline{x}$.

This implies further that G is an $n \times N$ matrix and $\underline{z}$ is an N-component vector, then

$$\int_{A+G\underline{z}} f(\underline{t})\, d\underline{t}$$

is a logarithmic concave function of the variable $\underline{z}$.

Consider now the random vector of components $\xi_1, \ldots, \xi_{n+N}$, denote by $\underline{e}$ its expectation vector and by C its covariance matrix. By assumption this random vector has a normal distribution. Assume also that this distribution is nondegenerated. Then the probability density of this random vector exists and it is given by

$$f(\underline{x}) = \left(\frac{\det C^{-1}}{(2\P)^n}\right)^{\frac{1}{2}} e^{-\frac{1}{2}(\underline{x}-\underline{e})' C^{-1}(\underline{x}-\underline{e})}, \quad \underline{x} \in R^{n+N}.$$

Out of the random vector components $\xi_1, \ldots, \xi_{n+N}$ we form two random vectors $\underline{\xi}^P$, $\underline{\xi}^F$ which are the following

(3.1)
$$\underline{\xi}^P = \begin{pmatrix} \xi_1 \\ \cdot \\ \cdot \\ \cdot \\ \xi_n \end{pmatrix}, \qquad \underline{\xi}^F = \begin{pmatrix} \xi_{n+1} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{n+N} \end{pmatrix}$$

and partition $\underline{e}$, $\underline{x}$ accordingly. The obtained parts will be denoted by $\underline{e}^P$, $\underline{e}^F$ resp. $\underline{x}^P$, $\underline{x}^F$.
Thus we have

$$(3.2) \qquad \underline{e}^P = E(\underline{\xi}^P) = \begin{pmatrix} e_1 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{pmatrix}, \quad \underline{e}^F = E(\underline{\xi}^F) = \begin{pmatrix} e_{n+1} \\ \cdot \\ \cdot \\ \cdot \\ e_{n+N} \end{pmatrix} .$$

The superscripts P and F are initials of the words "Past" resp. "Future". Let is rearrange and then partition the covariance matrix C so that we obtain the following

$$(3.3) \qquad \begin{pmatrix} c_{n+1,n+1} & \cdots & c_{n+1,n+N} & c_{n+1,1} & \cdots & c_{n+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n+N,n+1} & \cdots & c_{n+N,n+N} & c_{n+N,1} & \cdots & c_{n+N,n} \\ c_{1,n+1} & \cdots & c_{1,n+N} & c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n,n+1} & \cdots & c_{n,n+N} & c_{n1} & \cdots & c_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} E[(\underline{\xi}^F - \underline{e}^F)(\underline{\xi}^F - \underline{e}^F)'] & E[(\underline{\xi}^F - \underline{e}^F)(\underline{\xi}^P - \underline{e}^P)'] \\ E[(\underline{\xi}^P - \underline{e}^P)(\underline{\xi}^F - \underline{e}^F)'] & E[(\underline{\xi}^P - \underline{e}^P)(\underline{\xi}^P - \underline{e}^P)'] \end{pmatrix} =$$

$$= \begin{pmatrix} S & U \\ U' & T \end{pmatrix} .$$

It is known that the probability distribution of the random vector $\underline{\xi}^F$ given $\underline{\xi}^P = \underline{x}^P$ is a normal distribution with expectation vector

$$(3.4) \qquad \underline{e}^C = \underline{e}^F + UT^{-1}(\underline{x}^P - \underline{e}^P)$$

and covariance matrix

(3.5) $$S - UT^{-1}U' \quad ,$$

where the superscript in $\underline{e}^C$ refer to the word "Conditional". Thus the conditional probability density of $\underline{\xi}^F$ given $\underline{\xi}^P=\underline{x}^P$ is the following

$$f(\underline{x}^F|\underline{x}^P) = \left[\frac{\det(S-UT^{-1}U')}{(2\P)^N}\right]^{\frac{1}{2}} e^{-\frac{1}{2}(\underline{x}^F-\underline{e}^C)'(S-UT^{-1}U')(\underline{x}^F-\underline{e}^C)}.$$

The function $f(\underline{x}^F|\underline{x}^P)$ is logarithmic concave as a function of all variables in $\underline{x}^F$ and $\underline{x}^P$. Now we only need the fact it is logarithmic concave in $\underline{x}^F$ for every fixed $\underline{x}^P$. Consider the following set in the space of the vectors $\underline{x}^F$:

$$A = \{\underline{x}^F \mid a_k - \zeta_o - x_1 - \ldots - x_n + z_1 + \ldots + z_n \leq$$

$$\leq x_{n+1} + \ldots + x_{n+k} \leq$$

$$\leq b_k - \zeta_o - x_1 - \ldots - x_n + z_1 + \ldots + z_n, \quad k = 1,\ldots,N\}.$$

Then probability in the objective function of Problem (2.4) can be expressed in the following manner:

(3.8) $$P(a_k \leq \zeta_k - Z_k \leq b_k, \quad k=n+1,\ldots,n+N \mid \xi_1=x_1,\ldots,\xi_n=x_n) = \int_{A(z_{n+1},\ldots,z_{n+N})} f(\underline{x}|\underline{x}^P)\,dx,$$

where

(3.9) $$A(z_{n+1},\ldots,z_{n+N}) = A + \begin{pmatrix} z_{n+1} \\ z_{n+1}+z_{n+2} \\ \vdots \\ z_{n+1}+z_{n+2}+\ldots+z_{n+N} \end{pmatrix}.$$

Theorem 1 and Relation (3.8) jointly imply

Theorem 2. The probability standing in (3.8) is a logarithmic concave function of the variables $z_{n+1}, \ldots, z_{n+N}$.

In [1] the following theorem is proved.

Theorem 3. Let A be a convex set in R^m, symmetric about the origin. Let f be a quasi-concave probability density in R^m with the property that $f(-\underline{x}) = f(\underline{x})$ for every $\underline{x} \in R^m$. Then for every $\underline{y} \in R^m$ and $0 \leq k \leq 1$ we have the inequality

$$\int_{A+k\underline{y}} f(\underline{x})\, d\underline{x} \geq \int_{A+\underline{y}} f(\underline{x})\, d\underline{x} .$$

In other words, this theorem states that the probability of the set $A + t\underline{y}$ is a monotonically decreasing function in $[0,\infty]$ of the variable t for every fixed $\underline{y}$.

Theorem 3 implies that if we take unconstrained maximum of the probability (3.8) where only $z_{n+1}, \ldots, z_{n+N}$ are the variables, then the maximizing $z_{n+1}, \ldots, z_{n+N}$ satisfy the equalities

$$(3.10) \quad \frac{a_k + b_k}{2} - \zeta_{\text{initial}} - e^{C}_{n+1} - \ldots - e^{C}_{n+k+z_{n+1}+\ldots+z_{n+k}} = 0, \quad k=1,\ldots,N,$$

where

$$(3.11) \quad \zeta_{\text{initial}} = \zeta_o + x_1 + \ldots + x_n - z_1 - \ldots - z_n$$

is the water content of the lake beginning of Period n+1. This is that period for which we want to find an optimal policy.

4. SOLUTION OF THE PROBLEMS FORMULATED IN SECTION 2.

We shall consider Problem (2.4). Problem (2.3) is very similar and does not need a separate treatment. First we show that the optimization of the function (3.8) on the cube $0 \leq z_k \leq K$, $k = n+1,\ldots,n+N$ can be reduced to maximizations of the same functions on at most N faces of this cube. Sometimes the constrained optimal solution can be obtained directly without any computation. In fact first we solve the system of equations (3.10) with respect to $z_{n+1}, \ldots, z_{n+N}$. If we have $0 \leq z_k \leq K$ for $k = n+1,\ldots,n+N$, then this is the optimal solution also to the constrained problem (2.4). On the other hand, if for some i we have $z_i > K$ or for some k we have $z_k < 0$, then by Theorem 3 the optimum is attained on one of those faces of the cube which can be "seen" from the point with coordinates $z_{n+1}, \ldots, z_{n+N}$. These faces can be generated as follows. If $z_i > K$, then we adopt the face

$$z_i = K, \quad 0 \leq z_j \leq K, \quad j = n+1,\ldots,i-1,i+1,\ldots,n+N;$$

if $z_k < 0$, then we adopt the face

$$z_k = 0, \quad 0 \leq z_j \leq K, \quad j = n+1,\ldots,k-1,k+1,\ldots,n+N.$$

Obviously the number of such faces is at most N.

Example. Let $n = 2$, $N = 4$ and $z_3 > K$, $0 \leq z_4 \leq K$, $z_5 < 0$, $z_6 < K$. Then our face collection consists of the following three faces

$$\{ z_3,z_4,z_5,z_6 \mid z_3 = K, \quad 0 \leq z_4,z_5,z_6 \leq K \},$$
$$\{ z_3,z_4,z_5,z_6 \mid z_5 = 0, \quad 0 \leq z_3,z_4,z_6 \leq K \},$$
$$\{ z_3,z_4,z_5,z_6 \mid z_6 = K, \quad 0 \leq z_3,z_4,z_5 \leq K \}.$$

When optimizing in the faces we can apply various nonlinear programming methods. We have tested on this and similar stochastic programming problems the method of feasible directions, SUMT, GRG the flexible tolerance method and the cutting plane method.

It is worth to describe shortly the application of the SUMT interior point method. Let us assume that we want to maximize the function (*3.8*) on the following face

$$\{z_{n+1}, \ldots, z_{n+N} \mid z_{n+1}=K, \quad 0 \leq z_i \leq K, \quad i=n+2,\ldots,n+N\}.$$

If instead of the function (*3.8*) we work with its logarithm, then the penalty function is given by the following formula

(*4.1*)

$$\log P(a_k \leq \zeta_k - Z_k \leq b_k, \quad k = n+1,\ldots,n+N \mid \xi_1,\ldots,\xi_n) - r \sum_{k=n+2}^{n+N} \log z_k (1-z_k),$$

where r is a fixed positive number and $z_{n+1} = K$ in the sums $Z_k = z_1 + \ldots + z_k$, $k = n+1,\ldots,n+N$. The function (*4.1*) is convex and this fact makes the solutions of the unconstrained minimization problems relatively comfortable. Several unconstrained optimization method can be applied here [7] and we have tested a number of them. The use of a gradient free method seems to be advisable. The gradient of the function (*4.1*) is expressed in [10] but the formula is sophisticated.

Function values are computed by simulation at every step using a fast random number generation technique written in COMPASS for the CDC 3300 computer [3].

If we take a decreasing sequence $r_1, r_2, \ldots$ tending to zero, then the SUMT interior point method converges (the conditions are trivially satisfied in out case) in the sense that (*4.1*) converges to the negative logarithm of the optimum value of Problem (*2.4*).

If $N = 2$, then the original problem (2.4) is two-dimensional but since the faces of the rectangle

$$\{z_{n+1}, z_{n+2} \mid 0 \leq z_{n+1}, z_{n+2} \leq K\}$$

are lines we have to optimize on at most two lines. This is done by the use of the Fibonacci search [14].

5. METHOD FOR THE REGULATION OF THE WATER LEVEL OF LAKE BALATON

Having performed a large number of computations it turned out that using only two conditioning random variables (instead of the whole past history) and optimizing for two steps ahead i.e. choosing $N = 2$, a satisfactory water level control methodology can be obtained.

The lake is represented by a prism the surface of which is 600 km^2. We choose as water quantity unit that quantity which increases the water level by exactly 1 mm. (This quantity equals 600 000 m^3.) All date will be given in this unit.

According to what is said above, four random variables will be involved in every optimization problem. They belong to four consecutive months and will be denoted by ξ_1, ξ_2, ξ_3, ξ_4 in agreement with the earlier notations. To the earliest month corresponds ξ_1, then comes ξ_2 etc.

The prescribed lower resp. upper bounds are as follows:

	Lower bounds	Upper bounds
February- June	3100 mm	3400 mm
July-January	3000 mm	3300 mm

The originally prescribed levels communicated to us were 2900 mm resp. 3400 mm for every month. We observed that our control methodology allowed to keep the water level between the narrower limits 3000 mm resp. 3300 mm (with a satisfactory probability). However, due to large input water quantities in the first half of the year, the corresponding limits were increased by 100 mm and this improved the controllability for the most important summer months. Thus in all cases we have to solve the following type of problem:

$$\text{maximize} \quad P\left(\begin{array}{l} a_3 \leq \zeta_{\text{initial}} + \xi_3 - z_3 \leq b_3 \\ a_4 \leq \zeta_{\text{initial}} + \xi_3 + \xi_4 - z_3 - z_4 \leq b_4 \end{array} \middle| \, \xi_1, \xi_2 \right)$$

(5.1)

subject to

$$0 \leq z_3 \leq 200 ,$$

$$0 \leq z_4 \leq 200 ,$$

where a_3, b_3, a_4, b_4 are chosen according to the above table of the lower resp. upper bounds. The subscripts of a_3, b_3, z_3 and a_4, b_4, z_4 are chosen in accordance with the subscripts of ξ_3 and ξ_4.

It will be still more comfortable to operate with the following transformed random variables

(5.2) $$\zeta_1 = \xi_1, \quad \zeta_2 = \xi_2, \quad \zeta_3 = \xi_3, \quad \zeta_4 = \xi_3 + \xi_4.$$

The covariance matrix D of the random variables $\zeta_1, \zeta_2, \zeta_3, \zeta_4$ can be obtained from the covariance matrix

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{12} & c_{22} & c_{23} & c_{24} \\ c_{13} & c_{23} & c_{33} & c_{34} \\ c_{14} & c_{24} & c_{34} & c_{44} \end{pmatrix}$$

of the random variables $\xi_1, \xi_2, \xi_3, \xi_4$. We have

$$(5.3) \qquad D = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{13}+c_{14} \\ c_{12} & c_{22} & c_{23} & c_{23}+c_{24} \\ c_{13} & c_{23} & c_{33} & c_{33}+c_{34} \\ c_{13+14} & c_{23+24} & c_{33+34} & c_{33}+c_{44}+2c_{34} \end{pmatrix} .$$

With the aid of $\zeta_1,\zeta_2,\zeta_3,\zeta_4$, Problem (5.1) can be written in the following manner

$$\text{maximize} \quad P \left(\begin{matrix} a_3 \leq \zeta_{\text{initial}} + \zeta_3 - z_3 \leq b_3 \\ a_4 \leq \zeta_{\text{initial}} + \zeta_4 - z_3 - z_4 \leq b_4 \end{matrix} \;\middle|\; \zeta_1, \zeta_2 \right)$$

(5.4)

$$\text{subject to} \qquad 0 \leq z_3 \leq 200 ,$$
$$0 \leq z_4 \leq 200 .$$

Let us rearrange and then partition the covariance matrix D in a way indicated here below

(5.5)

ζ_3	ζ_4	ζ_1	ζ_2	
S_1		U_1		ζ_3
				ζ_4
U_1'		T		ζ_1
				ζ_2

Since $\zeta_1 = \xi_1$, $\zeta_2 = \xi_2$ it follows that $T_1 = T$ where T is taken out from that special case of (3.3) in which $n = 2$, $N = 2$. Then we have

$$(5.6)\quad E\left[\begin{pmatrix}\zeta_3\\ \zeta_4\end{pmatrix}\middle|\begin{pmatrix}\zeta_1\\ \zeta_2\end{pmatrix}\right]=\begin{pmatrix}e_3\\ e_3+e_4\end{pmatrix}+U_1T_1^1\left[\begin{pmatrix}\zeta_1\\ \zeta_2\end{pmatrix}-\begin{pmatrix}e_1\\ e_2\end{pmatrix}\right].$$

On the next pages we present the input data (Table I.) for the 50 years between 1921 and 1970. Taking into account a longer (less reliable but improved by hydrological considerations) time series, the Institute of Water Management of the Technical University of Budapest advised the use of a Gaussian process as the mathematical model of the water input process, after a careful statistical analysis [2].

Together with the realized values of the input process we present the corresponding expectations and variances.

Then we present a part of a 24×24 correlation matrix (Table II.)). Assuming that the random input of one month is stochastically independent of the inputs of such months which are farther then one year then it turns out that all nonzero correlation will be contained in a 24×24 correlation matrix. In practice, however, only correlations very near the diagonal will be needed because the others are very small. That part of the correlation matrix what we present here is larger than the necessary part but is well illustrates that dependencies exist only between very near months.

As it was mentioned in the Introduction we carried out the monthly optimizations between 1922 and 1970.

First we computed twelwe D matrices and to each D we also computed the corresponding two matrices $U_1 T_1^{-1}$, $S_1 - U_1 T_1^{-1} U_1'$. These matrices are fixed, they do not depend on actual values of the input time series. Then using the actual values of ζ_1 and ζ_2 we computed all conditional expectations (*5.6*). Finally came the 588 optimizations (one for every month in the years 1922-1970) out of which a large number were trivial i.e., the solution of the equation (*3.10*) specialized to our case (n = 2, N = 2) produced such z_3 and z_4 for which $0 \leq z_3 , z_4 \leq 200$.

As an example we consider the problem of finding the optimal water quantity to be released in July, 1953. In this case the random variables $\xi_1, \xi_2, \xi_3, \xi_4$ have the following meanings

ξ_1 input water quantity in May, 1953

ξ_2 input water quantity in June 1953

ξ_3 input water quantity in July 1953

ξ_4 input water quantity in Aug. 1953

The expectations, dispersions and the correlation matrix can be obtained from the presented tables. They are reproduced here:

$E(\xi_1)=29.78$, $E(\xi_2)=14.52$, $E(\xi_3)=-43.44$, $E(\xi_4)=-38.30$,

$D(\xi_1)=63.11$, $D(\xi_2)=73.98$, $D(\xi_3)= 73.96$, $D(\xi_4)= 69.58$,

$$R = \begin{array}{c} \begin{matrix} \xi_1 & \xi_2 & \xi_3 & \xi_4 \end{matrix} \\ \begin{pmatrix} 1.000 & 0.333 & 0.198 & 0.201 \\ 0.333 & 1.000 & 0.579 & 0.263 \\ 0.198 & 0.579 & 1.000 & 0.352 \\ 0.201 & 0.263 & 0.352 & 1.00 \end{pmatrix} \end{array} \begin{matrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{matrix}$$

The transformed variables (5.2) have the following expectations and covariance matrix

$$E(\zeta_1) = 27.98, \quad E(\zeta_2) = -4.52 , \quad E(\zeta_3) = -43.44, \quad E(\zeta_4) = -81.74,$$

$$D = \begin{array}{c} \begin{matrix} \zeta_1 & \zeta_2 & \zeta_3 & \zeta_4 \end{matrix} \\ \begin{pmatrix} 3982.87210011 & 1554.73630744 & 924.18788880 & 1806.81784260 \\ 1554.73630744 & 5473.04040002 & 3168.03370320 & 4521.83367240 \\ 924.18788880 & 3168.03370320 & 5470.08160018 & 7281.52175378 \\ 1806.81784260 & 4521.83367240 & 7281.52175378 & 13934.33830761 \end{pmatrix} \end{array} \begin{matrix} \\ \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{matrix}$$

From here we obtain

$$S_1 = \begin{pmatrix} 5470.08160018 & 7281.52175378 \\ 7281.52175378 & 13934.33830761 \end{pmatrix} ,$$

$$U_1 = \begin{pmatrix} 924.18788880 & 3168.03370320 \\ 1806.81784260 & 4521.83367240 \end{pmatrix} ,$$

$$T_1 = \begin{pmatrix} 3982.87210011 & 1554.73630744 \\ 1554.73630744 & 5473.04040002 \end{pmatrix} ,$$

$$T_1^{-1} = \begin{pmatrix} 0.00028239 & -0.00008022 \\ -0.00008022 & 0.00020550 \end{pmatrix} ,$$

$$U_1 T_1^{-1} = \begin{pmatrix} 0.00684480 & 0.57689906 \\ 0.14748962 & 0.78430377 \end{pmatrix} .$$

The realised water input data are the following

$$\xi_1 = \zeta_1 = 40 , \qquad \xi_2 = \zeta_2 = 22,$$

hence the conditional expectation equals

$$E\left[\begin{pmatrix}\zeta_3\\ \zeta_4\end{pmatrix}\,\middle|\,\zeta_1 = 40, \quad \zeta_2 = 22\right] =$$

$$(5.7) \qquad = \begin{pmatrix}-43.44\\ -81.74\end{pmatrix} + U_1 T_1^{-1}\left[\begin{pmatrix}40\\ 22\end{pmatrix} - \begin{pmatrix}29.78\\ -4.52\end{pmatrix}\right] = \begin{pmatrix}-28.07\\ -59.43\end{pmatrix}.$$

The covariance matrix of the conditional distribution of ζ_3 , ζ_4 given that $\zeta_i = 40$, $\zeta_2 = 22$ is the following

$$(5.8) \qquad S_1 - U_1 \bar{T}_1^{1} U_1' = \begin{pmatrix}3636.12006366 & 4660.51286423\\ 4660.51286423 & 10121.36024427\end{pmatrix}.$$

Since $\zeta_{\text{initial}} = 3205$, the optimization problem (5.4) can be written in the following manner

$$(5.9) \qquad \begin{aligned} &\text{maximize} \quad P\left(\begin{matrix}-2.05 \leqq \zeta_3 - z_3 \leqq 95\\ -2.05 \leqq \zeta_4 - z_3 - z_4 \leqq 95\end{matrix}\,\middle|\,\zeta_1 = 40, \quad \zeta_2 = 22\right)\\ &\text{subject to}\\ &\qquad 0 \leqq z_3 \leqq 200,\\ &\qquad 0 \leqq z_4 \leqq 200, \end{aligned}$$

and the above probability distribution is two-dimensional normal with expectation vector (5.7) and covariance matrix (5.8)

If we compute z_3, z_4 according to (3.10), we obtain the values

$$z_3 = 27, \quad z_4 = -31.$$

It follows that the optimal z_4 to Problem (5.9) equals

$$z_4^* = 0$$

and z_3^* is the optimal solution of the following one-dimensional problem

$$\text{maximize} \quad P\left(\begin{matrix} -205 \leq \zeta_3 - z_3 \leq 95 \\ -205 \leq \zeta_4 - z_3 \leq 95 \end{matrix} \;\middle|\; \zeta_1 = 40, \quad \zeta_2 = 22 \right), \tag{5.10}$$

subject to

$$0 \leq z_3 \leq 200.$$

The Fibonacci search gives in 15 steps the result

$$z_3^* = 2.$$

In higher dimensional cases the values of the objective function are determined by simulation. In the two-dimensional case numerical integration is satisfactorily effective. We use a reduction formula and then one-dimension numerical integration. The reduction formula states that if $\phi(x,y;r)$ and $\phi(x,y;r)$ are the two-dimensional normal probability density resp. distribution functions with standard marginal distributions and r is the correlation coefficient ($|r| < 1$), then we have

$$\int_a^b \int_c^d \phi(x,y;r)\, dydx = \int_a^b \left[\phi\left(\frac{d-rx}{\sqrt{1-r^2}}\right) - \phi\left(\frac{c-rx}{\sqrt{1-r^2}}\right) \right] \rho(x)\, dx$$

where ϕ and ϕ denote the one-dimensional standard normal density resp. distribution function. The one-dimensional numerical integration is done by the Romberg-Havie procedure [4]. The computational precision is 10^{-3}.

In Table III. we summarize the results of the 588 optimizations. The + resp. - signs mean that the water level is higher resp. lower than desired. We see that only a few such signs occur. Moreover the first month of 1922 do not count because we need a few periods for the running-in of this control methodology. In 1946 the lock was repaired

which caused a high water level. The variation of the controlled water level is illustrated on *Fig. 2*.

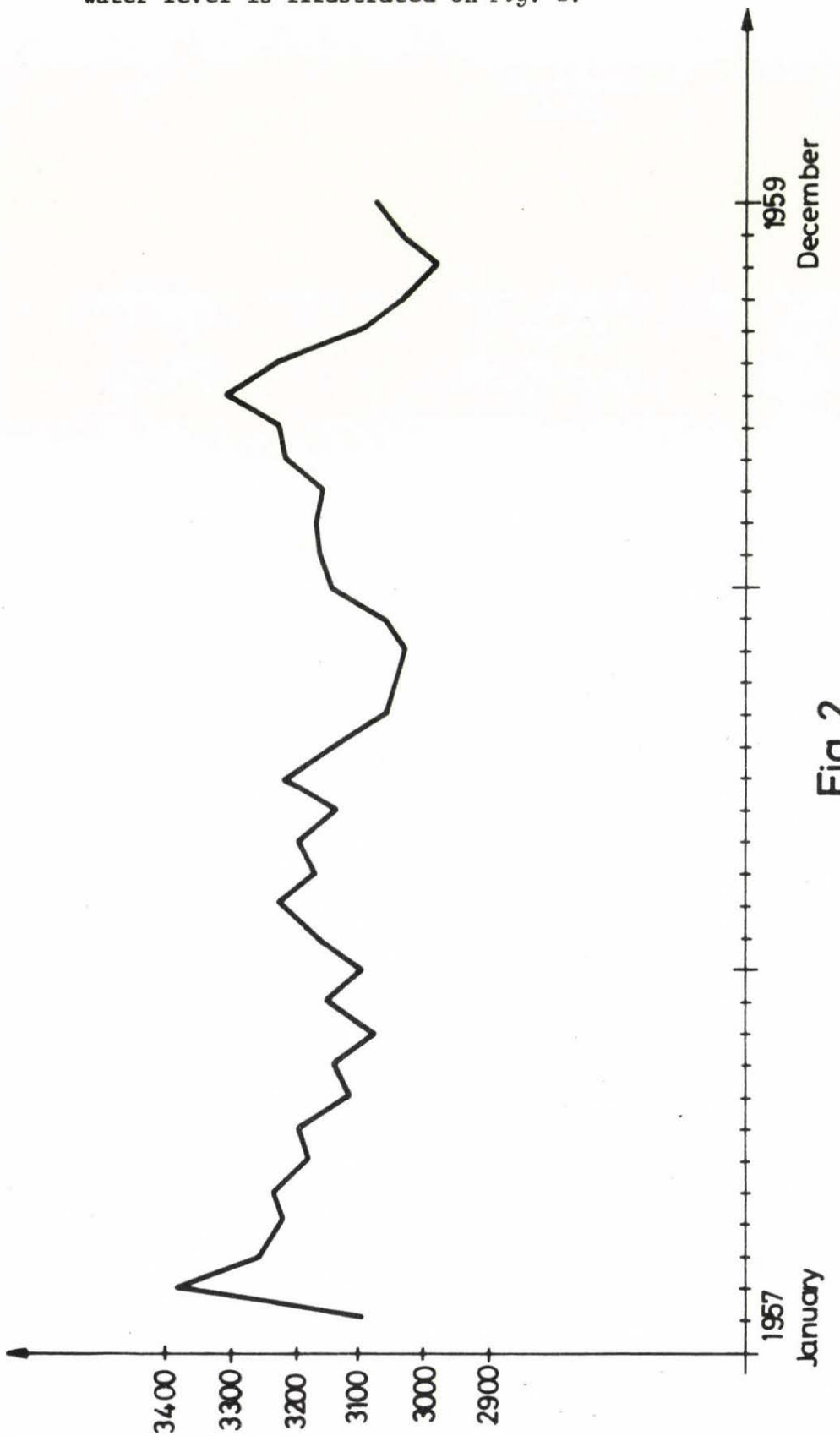


Fig. 2
Variation of the controlled water level of Lake Balaton (Illustration for three years)

REFERENCES

[1] Anderson,T.W., "The Integral of a Symmetric Unimodal Function over a Symmmetric Convex Set and Some Probability Inequalities" Proc.Amer.Math.Soc. 6(1955), 1970-1976.

[2] The Balaton (Institute for Water Management of the Technical University of Budapest, Budapest, 1970, in Hungarian).

[3] Deák I., "A Fast Monte Carlo Method for Computing Probabilities of Sets in Higher Dimensional Spaces in Case of Normal Distribution", Water Resources Research to appear.

[4] Havie,T., "On a modification of Romberg's algorithm", BIT 6(1966) 24-30.

[5] Himmelblau, D.M., Applied Nonlinear Programming (McGraw-Hill, New York, 1972).

[6] Fiacco, A.W. and McCormick, G.P., Nonlinear Programming Sequential Unconstrained Minimization Techniques (Wiley, New York, London 1968).

[7] Powell, M.J.D., "Recent Advances in Unconstrained Optimization". Mathematical Programming 1(1971) 26-57.

[8] Prékopa,A., "On Optimization of Stochastic Systems", Academic Doctorate Theses, Budapest, 1970.

[9] Prékopa,A., "Logarithmic Concave Measures with Application to Stochastic Programming", Acta Scientiarum Mathematicarum 32(1971) 301-316.

[10] Prékopa A., "Optimal Control of a Storage Level Using Stochastic Programming", Problems of Control and Information Theory 4(1975) 193-204.

[11] Prékopa,A., and Szántai,T., "On Multi-Storage Stochastic Programming (with Application to Optimal Control of Water Supply)", in. Colloquia Mathematica Societatis János Bolyai 12. Progress in Operations Research Ed. A. Prékopa (János Bolyai Mathematical Society Budapest and North Holland Publishing Company Amsterdam, London, 1976) 733-755.

[12] Szesztai,K., The Hydrology of Lake Balaton (Institute for Water Management, Budapest, 1962, in Hungarian).

[13] Wilks,S.S., Mathematical Statistics, (Wiley, New York, London 1962).

[14] Wilde,D.J., Optimum Seeking Methods, (Prentice Hall Inc.,Englewood Cliffs, N.J. 1964).

Table I.

Natural Water Content Changes of Lake Balaton (Rainfall+Inflow-Evaporation)

	January	February	March	April	May	June	July	August	September	October	November	December
1921	62	133	16	26	43	-35	-102	-143	-68	4	77	58
1922	102	119	128	192	-3	-52	-103	-18	118	198	96	72
1923	85	122	239	94	-17	9	-52	-75	-30	60	119	154
1924	38	94	262	171	104	26	-50	-7	-4	-26	10	56
1925	48	75	86	48	54	-18	-22	-56	87	-8	232	106
1926	130	165	77	23	-7	74	101	97	-11	89	173	139
1927	121	79	137	62	-25	-48	-54	32	45	-1	19	36
1928	99	126	111	31	90	-51	-97	-59	81	49	76	73
1929	95	74	170	208	38	-24	-90	-101	-76	39	159	16
1930	79	131	106	172	-51	-95	-124	-27	-21	182	178	235
1931	178	208	302	166	37	-59	-156	-32	43	16	92	37
1932	101	35	135	65	59	-95	-97	-65	-59	74	18	48
1933	51	73	92	21	41	-1	-119	-41	-8	48	240	136
1934	137	69	77	-37	-62	-19	-42	-46	27	9	103	72
1935	63	147	85	23	-21	-85	-136	-57	-13	18	31	172
1936	162	199	107	63	85	4	-90	-107	27	126	97	94
1937	105	146	343	260	-4	40	8	10	67	99	222	331
1938	214	118	71	18	79	-39	-98	58	-11	34	10	87
1939	114	77	73	-61	74	9	-137	-37	2	97	92	77
1940	68	67	339	111	96	76	22	159	247	174	243	90
1941	127	231	204	216	129	-39	-83	-35	-57	70	200	148
1942	115	215	367	307	178	-66	-57	-99	-56	-27	23	61
1943	112	179	-12	-5	-13	108	32	-106	-2	-2	148	129
1944	57	113	241	11	52	96	-32	-72	-50	146	225	254
1945	191	309	215	35	-37	-79	-68	-99	-28	19	99	97
1946	68	118	60	-41	-75	-51	-63	-93	-52	7	86	84
1947	87	182	448	117	-17	-54	-77	-148	-85	-5	22	88
1948	110	106	8	85	-19	17	136	-41	-30	46	47	137
1949	85	9	-16	4	41	-73	-77	-75	-84	-2	175	52
1950	115	188	63	63	-33	-155	-90	-58	-1	67	216	190
1951	139	154	206	-27	112	167	39	-35	18	-36	21	85
1952	114	155	166	37	-49	-38	-153	-128	-26	133	97	155
1953	110	71	18	-5	40	22	-93	-51	-59	8	-22	17
1954	73	56	192	28	175	16	22	-54	-31	21	76	124
1955	121	134	152	66	10	-46	39	81	26	140	185	77
1956	88	108	160	114	63	14	-3	-75	-85	22	54	114
1957	46	293	75	17	12	-56	19	-68	14	-26	78	36
1958	99	89	47	24	-60	89	-29	-83	-5	-19	32	85
1959	98	40	11	59	9	82	63	-55	-69	-40	46	140
1960	127	147	67	63	33	-80	40	-57	-7	141	154	207
1961	121	113	39	13	50	8	-80	-110	-70	-11	90	51
1962	116	70	185	95	-20	-50	21	-115	-42	-1	220	149
1963	191	130	368	174	9	-17	-127	7	96	75	71	109
1964	42	73	250	123	77	-2	-30	-31	-2	170	82	190
1965	164	111	131	190	193	244	145	131	72	15	187	448
1966	122	202	151	146	46	38	89	87	48	19	240	203
1967	204	184	143	141	-11	96	-55	-107	60	17	33	40
1968	104	85	50	10	-65	-100	-140	70	21	40	167	69
1969	139	313	234	48	14	84	-58	-30	11	20	72	110
1970	111	182	382	223	35	-18	-64	49	-5	12	62	85
Expectations	108.96	132.34	151.22	79.74	29.78	-4.52	-43.44	-38.30	-0.74	46.00	109.46	114.46
Dispersions	41.52	67.04	112.84	83.51	63.11	73.98	73.96	69.58	61.95	62.51	75.15	80.60

Table II.

Correlations of Natural Water Content Changes of Lake Balaton

	Jan.	Febr.	March.	April.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Febr.	March.	April.	May.	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Jan.	1,00	0,36	0,07	0,11	-0,01	0,13	-0,02	0,15	0,16	-0,05	-0,17	-0,02	-0,18											
Febr.	0,36	1,00	0,28	0,20	-0,06	-0,05	-0,02	-0,10	0,04	-0,07	-0,04	0,05	0,02	-0,00										
March.	0,07	0,28	1,00	0,63	0,24	0,04	-0,09	0,14	0,22	0,13	-0,02	0,12	0,08	0,09	-0,02									
April.	0,11	0,20	0,63	1,00	0,28	-0,02	0,05	0,19	0,17	0,16	0,11	0,25	0,16	0,08	-0,12	0,23								
May.	-0,01	-0,06	0,24	0,28	1,00	0,33	0,20	0,20	0,07	-0,05	-0,03	0,14	0,01	0,11	-0,03	0,23	0,14							
June	0,13	-0,05	0,04	-0,02	0,33	1,00	0,58	0,26	0,22	-0,15	0,02	0,34	0,03	0,04	0,02	0,11	-0,07	-0,08						
July	-0,02	-0,02	-0,09	0,05	0,20	0,58	1,00	0,35	0,12	-0,12	0,22	0,36	0,17	0,05	-0,19	-0,03	-0,09	0,02	0,17					
Aug.	0,15	-0,10	0,14	0,19	0,20	0,26	0,35	1,00	0,60	0,26	0,30	0,29	0,19	0,20	0,08	0,10	0,31	0,08	-0,03	0,15				
Sept.	0,16	0,04	0,22	0,17	0,07	0,22	0,12	0,60	1,00	0,35	0,28	0,13	0,09	0,02	0,10	0,15	0,17	0,15	-0,08	0,21	-0,05			
Oct.	-0,05	-0,07	0,15	0,16	-0,05	-0,15	-0,12	0,26	0,35	1,00	0,39	0,31	0,20	0,11	0,26	0,27	0,32	0,13	-0,09	0,18	-0,02	-0,09		
Nov.	-0,17	-0,04	-0,02	0,11	-0,03	0,02	0,22	0,30	0,28	0,39	1,00	0,51	0,53	0,39	0,26	0,22	-0,04	0,04	-0,05	0,14	-0,00	-0,11	-0,11	
Dec.	-0,02	0,05	0,12	0,25	0,14	0,34	0,36	0,29	0,13	0,31	0,51	1,00	0,56	0,55	0,11	0,10	0,19	0,20	0,16	0,26	0,12	-0,17	-0,10	0,04
Jan.	-0,18	0,02	0,08	0,16	0,01	0,03	0,17	0,19	0,09	0,20	0,59	0,56	1,00	0,56	0,07	0,11	-0,01	0,13	-0,02	0,15	0,16	-0,05	-0,17	-0,02
Febr.		-0,00	0,09	0,08	0,11	0,04	0,05	0,20	0,02	0,11	0,39	0,35	0,36	1,00	0,28	0,20	-0,06	-0,05	-0,02	-0,10	0,04	-0,07	-0,04	0,05
March.			-0,02	-0,12	-0,03	0,02	-0,19	0,08	0,10	0,26	0,26	0,11	0,07	0,28	1,00	0,63	0,24	0,04	-0,09	0,14	0,22	0,15	-0,03	0,12
April.				0,23	0,23	0,11	-0,05	0,10	0,15	0,27	0,22	0,10	0,11	0,20	0,63	1,00	0,28	-0,02	0,05	0,19	0,17	0,16	0,11	0,25
May.					0,14	-0,07	-0,09	0,31	0,17	0,32	-0,04	0,19	-0,01	-0,06	0,24	0,28	1,00	0,33	0,20	0,20	0,07	-0,05	-0,03	0,14
June						-0,08	0,02	0,08	0,15	0,13	0,04	0,20	0,13	-0,05	0,04	-0,02	0,33	1,00	0,58	0,26	0,22	-0,15	0,02	0,34
July							0,17	-0,03	-0,08	-0,09	-0,05	0,16	-0,02	-0,02	-0,09	0,05	0,20	0,58	1,00	0,35	0,12	-0,12	0,22	0,36
Aug.								0,15	0,21	0,18	0,14	0,26	0,15	-0,10	0,14	0,19	0,20	0,26	0,35	1,00	**0,60**	0,26	0,30	0,29
Sept.									-0,05	-0,02	-0,00	0,12	0,16	0,04	0,22	0,17	0,07	0,22	0,12	0,60	1,00	0,35	0,28	0,13
Oct.										-0,09	-0,11	-0,17	-0,05	-0,07	0,13	0,16	-0,05	-0,15	-0,12	0,26	0,35	1,00	0,39	0,31
Nov.											-0,11	-0,10	-0,17	-0,04	-0,03	0,11	-0,03	0,02	0,22	0,30	0,28	0,39	1,00	0,51
Dec.												0,04	-0,02	0,05	0,12	0,25	0,14	0,34	0,36	0,29	0,13	0,31	0,51	1,00

Table III.

Numerical Results. Application of the Proposed Control Methodology for Lake Balaton between the years 1922-1970. The z_3 are the Final Accepted Values. Water Levels are Computed with These. + Sign (resp.-sign) Means Water Level Above 3400 mm (resp.Below 2900 mm).

		Z3	Z4	WATER LEVEL	PROBABILITY
1922	JANUARY	0.	0.	2314. (-)	0.00
1922	FEBRUARY	0.	0.	2433. (-)	0.00
1922	MARCH	0.	0.	2561. (-)	0.00
1922	APRIL	0.	0.	2753. (-)	0.00
1922	MAY	0.	0.	2750. (-)	0.00
1922	JUNE	0.	0.	2698. (-)	0.00
1922	JULY	0.	0.	2595. (-)	0.00
1922	AUGUST	0.	0.	2577. (-)	0.00
1922	SEPTEMBER	0.	0.	2695. (-)	0.00
1922	OCTOBER	0.	0.	2893. (-)	0.00
1922	NOVEMBER	0.	129.	2989.	74.10
1922	DECEMBER	0.	84.	3061.	90.06
1923	JANUARY	9.	22.	3137.	96.20
1923	FEBRUARY	3.	144.	3256.	70.79
1923	MARCH	155.	75.	3341.	61.29
1923	APRIL	200.	54.	3235.	88.62
1923	MAY	21.	0.	3196.	82.43
1923	JUNE	0.	0.	3205.	71.23
1923	JULY	0.	0.	3153.	84.81
1923	AUGUST	0.	0.	3078.	79.77
1923	SEPTEMBER	0.	0.	3048.	80.39
1923	OCTOBER	0.	40.	3108.	80.80
1923	NOVEMBER	67.	119.	3160.	77.34
1923	DECEMBER	131.	112.	3182.	90.72
1924	JANUARY	151.	41.	3069.	96.20
1924	FEBRUARY	0.	99.	3163.	67.16
1924	MARCH	52.	65.	3373.	61.29
1924	APRIL	200.	95.	3344.	85.70
1924	MAY	141.	0.	3307.	82.43
1924	JUNE	74.	73.	3258.	76.86
1924	JULY	66.	0.	3143.	85.98
1924	AUGUST	0.	0.	3136.	79.03
1924	SEPTEMBER	3.	57.	3129.	92.30
1924	OCTOBER	26.	116.	3077.	84.40
1924	NOVEMBER	7.	87.	3080.	77.34
1924	DECEMBER	0.	67.	3136.	90.50
1925	JANUARY	62.	0.	3122.	96.20
1925	FEBRUARY	0.	115.	3197.	70.34
1925	MARCH	76.	62.	3207.	61.29
1925	APRIL	5.	27.	3250.	88.72
1925	MAY	16.	0.	3288.	82.42
1925	JUNE	47.	63.	3223.	76.86
1925	JULY	0.	0.	3201.	85.30
1925	AUGUST	14.	0.	3131.	81.95
1925	SEPTEMBER	0.	6.	3218.	92.04
1925	OCTOBER	138.	121.	3071.	84.40
1925	NOVEMBER	26.	97.	3277.	77.34
1925	DECEMBER	200.	195.	3183.	77.45
1926	JANUARY	162.	63.	3151.	96.20
1926	FEBRUARY	40.	151.	3276.	70.79
1926	MARCH	192.	89.	3162.	61.29
1926	APRIL	0.	0.	3185.	86.00
1926	MAY	0.	0.	3178.	76.80
1926	JUNE	0.	0.	3252.	69.81
1926	JULY	81.	0.	3271.	85.83
1926	AUGUST	130.	13.	3238.	81.95
1926	SEPTEMBER	151.	58.	3076.	92.30
1926	OCTOBER	0.	119.	3165.	84.23
1926	NOVEMBER	139.	131.	3198.	77.34
1926	DECEMBER	200.	128.	3137.	90.72

		Z3	Z4	WATER LEVEL	PROBABILITY
1927	JANUARY	113.	53.	3146.	96.20
1927	FEBRUARY	37.	155.	3188.	70.79
1927	MARCH	61.	68.	3264.	61.29
1927	APRIL	85.	35.	3240.	88.72
1927	MAY	12.	0.	3203.	82.43
1927	JUNE	0.	0.	3155.	72.52
1927	JULY	0.	0.	3101.	61.25
1927	AUGUST	0.	0.	3133.	63.79
1927	SEPTEMBER	23.	70.	3155.	92.30
1927	OCTOBER	70.	134.	3084.	84.40
1927	NOVEMBER	34.	98.	3069.	77.34
1927	DECEMBER	0.	67.	3105.	90.44
1928	JANUARY	28.	0.	3176.	96.20
1928	FEBRUARY	40.	140.	3262.	70.79
1928	MARCH	161.	77.	3212.	61.29
1928	APRIL	23.	24.	3220.	88.72
1928	MAY	0.	0.	3310.	82.05
1928	JUNE	87.	71.	3172.	76.86
1928	JULY	0.	0.	3075.	72.15
1928	AUGUST	0.	0.	3016.	48.87
1928	SEPTEMBER	0.	0.	3097.	52.93
1928	OCTOBER	16.	120.	3130.	84.40
1928	NOVEMBER	107.	119.	3099.	77.34
1928	DECEMBER	48.	99.	3125.	90.72
1929	JANUARY	69.	17.	3150.	96.20
1929	FEBRUARY	20.	145.	3204.	70.79
1929	MARCH	78.	65.	3297.	61.29
1929	APRIL	133.	41.	3461.(+)	88.72
1929	MAY	200.	45.	3299.	76.92
1929	JUNE	25.	57.	3250.	76.86
1929	JULY	15.	0.	3146.	85.27
1929	AUGUST	0.	0.	3045.	75.32
1929	SEPTEMBER	0.	0.	2969.	54.15
1929	OCTOBER	0.	0.	3008.	38.95
1929	NOVEMBER	0.	69.	3167.	75.62
1929	DECEMBER	155.	123.	3028.	90.72
1930	JANUARY	0.	5.	3107.	96.19
1930	FEBRUARY	0.	101.	3238.	69.87
1930	MARCH	142.	77.	3202.	61.29
1930	APRIL	11.	23.	3363.	88.72
1930	MAY	145.	0.	3167.	82.38
1930	JUNE	0.	0.	3072.	53.06
1930	JULY	0.	0.	2948.	15.01
1930	AUGUST	0.	0.	2921.	2.97
1930	SEPTEMBER	0.	0.	2900.	8.62
1930	OCTOBER	0.	24.	3082.	14.85
1930	NOVEMBER	92.	166.	3168.	77.34
1930	DECEMBER	189.	129.	3214.	90.72
1931	JANUARY	200.	80.	3192.	96.20
1931	FEBRUARY	122.	171.	3278.	70.79
1931	MARCH	200.	108.	3380.	61.21
1931	APRIL	200.	103.	3346.	79.80
1931	MAY	147.	0.	3235.	82.43
1931	JUNE	0.	36.	3176.	76.62
1931	JULY	0.	0.	3020.	69.09
1931	AUGUST	0.	0.	2988.	17.73
1931	SEPTEMBER	0.	0.	3031.	48.74
1931	OCTOBER	0.	69.	3047.	81.97
1931	NOVEMBER	4.	105.	3136.	77.34
1931	DECEMBER	87.	104.	3086.	90.72

		Z3	Z4	WATER LEVEL	PROBABILITY
1932	JANUARY	26.	15.	3161.	96.20
1932	FEBRUARY	26.	140.	3170.	70.79
1932	MARCH	23.	57.	3282.	61.29
1932	APRIL	101.	40.	3245.	88.72
1932	MAY	18.	0.	3287.	82.43
1932	JUNE	47.	64.	3145.	76.86
1932	JULY	0.	0.	3048.	47.34
1932	AUGUST	0.	0.	2983.	33.58
1932	SEPTEMBER	0.	0.	2924.	26.62
1932	OCTOBER	0.	8.	2998.	19.06
1932	NOVEMBER	0.	88.	3016.	76.22
1932	DECEMBER	0.	44.	3064.	86.50
1933	JANUARY	0.	0.	3115.	96.05
1933	FEBRUARY	0.	108.	3188.	70.05
1933	MARCH	66.	51.	3215.	61.29
1933	APRIL	15.	28.	3220.	88.72
1933	MAY	0.	0.	3261.	81.88
1933	JUNE	18.	60.	3242.	76.86
1933	JULY	23.	0.	3101.	85.52
1933	AUGUST	0.	0.	3060.	60.82
1933	SEPTEMBER	0.	0.	3052.	85.17
1933	OCTOBER	0.	60.	3100.	82.45
1933	NOVEMBER	58.	115.	3281.	77.34
1933	DECEMBER	200.	200.	3217.	70.14
1934	JANUARY	200.	74.	3154.	96.20
1934	FEBRUARY	51.	155.	3172.	70.79
1934	MARCH	39.	67.	3210.	61.29
1934	APRIL	4.	27.	3169.	88.72
1934	MAY	0.	0.	3107.	71.05
1934	JUNE	0.	0.	3088.	33.21
1934	JULY	0.	0.	3046.	40.89
1934	AUGUST	0.	0.	3000.	45.18
1934	SEPTEMBER	0.	0.	3027.	42.11
1934	OCTOBER	0.	56.	3036.	80.52
1934	NOVEMBER	0.	90.	3139.	77.25
1934	DECEMBER	94.	107.	3117.	90.72
1935	JANUARY	66.	24.	3114.	96.20
1935	FEBRUARY	0.	117.	3261.	70.35
1935	MARCH	174.	80.	3172.	61.29
1935	APRIL	0.	0.	3195.	87.87
1935	MAY	0.	0.	3174.	78.89
1935	JUNE	0.	0.	3089.	57.04
1935	JULY	0.	0.	2953.	23.66
1935	AUGUST	0.	0.	2896.(-)	3.21
1935	SEPTEMBER	0.	0.	2883.(-)	1.42
1935	OCTOBER	0.	20.	2901.	9.20
1935	NOVEMBER	0.	35.	2932.	43.31
1935	DECEMBER	0.	24.	3104.	50.99
1936	JANUARY	60.	22.	3205.	96.20
1936	FEBRUARY	118.	161.	3286.	70.79
1936	MARCH	200.	108.	3193.	61.11
1936	APRIL	4.	14.	3252.	88.72
1936	MAY	21.	0.	3315.	82.42
1936	JUNE	86.	70.	3233.	76.86
1936	JULY	22.	0.	3121.	85.78
1936	AUGUST	0.	0.	3014.	70.77
1936	SEPTEMBER	0.	0.	3041.	29.59
1936	OCTOBER	0.	48.	3167.	81.86
1936	NOVEMBER	164.	146.	3100.	77.34
1936	DECEMBER	72.	105.	3122.	90.72

		Z3	Z4	WATER LEVEL	PROBABILITY
1937	JANUARY	75.	26.	3152.	96.20
1937	FEBRUARY	30.	148.	3269.	70.79
1937	MARCH	177.	83.	3435.(+)	61.29
1937	APRIL	200.	147.	3495.(+)	46.19
1937	MAY	200.	70.	3291.	64.33
1937	JUNE	2.	47.	3328.	76.86
1937	JULY	134.	0.	3203.	85.61
1937	AUGUST	32.	7.	3180.	81.95
1937	SEPTEMBER	52.	50.	3195.	92.30
1937	OCTOBER	115.	133.	3179.	84.40
1937	NOVEMBER	173.	137.	3228.	77.34
1937	DECEMBER	200.	180.	3359.	86.97
1938	JANUARY	200.	195.	3373.	15.74
1938	FEBRUARY	200.	200.	3291.	42.93
1938	MARCH	175.	85.	3187.	61.29
1938	APRIL	0.	2.	3205.	88.36
1938	MAY	0.	0.	3284.	80.17
1938	JUNE	58.	69.	3188.	76.86
1938	JULY	0.	0.	3090.	78.52
1938	AUGUST	0.	0.	3148.	55.80
1938	SEPTEMBER	56.	87.	3081.	92.30
1938	OCTOBER	0.	112.	3115.	84.26
1938	NOVEMBER	67.	109.	3057.	77.34
1938	DECEMBER	0.	58.	3144.	89.97
1939	JANUARY	80.	2.	3178.	96.20
1939	FEBRUARY	58.	148.	3198.	70.79
1939	MARCH	71.	67.	3200.	61.29
1939	APRIL	0.	18.	3139.	88.67
1939	MAY	0.	0.	3213.	58.54
1939	JUNE	0.	62.	3222.	76.84
1939	JULY	13.	0.	3072.	85.77
1939	AUGUST	0.	0.	3035.	46.82
1939	SEPTEMBER	0.	0.	3037.	76.64
1939	OCTOBER	0.	55.	3134.	80.96
1939	NOVEMBER	114.	134.	3112.	77.34
1939	DECEMBER	76.	104.	3113.	90.72
1940	JANUARY	61.	22.	3120.	96.20
1940	FEBRUARY	0.	126.	3187.	70.59
1940	MARCH	59.	61.	3466.(+)	61.29
1940	APRIL	200.	163.	3377.	30.64
1940	MAY	174.	3.	3300.	82.43
1940	JUNE	70.	72.	3305.	76.86
1940	JULY	149.	0.	3179.	86.13
1940	AUGUST	15.	14.	3322.	81.95
1940	SEPTEMBER	200.	160.	3369.	87.77
1940	OCTOBER	200.	200.	3343.	36.85
1940	NOVEMBER	200.	200.	3386.	20.00
1940	DECEMBER	200.	198.	3276.	9.88
1941	JANUARY	200.	117.	3203.	96.10
1941	FEBRUARY	88.	148.	3346.	70.79
1941	MARCH	200.	137.	3350.	53.40
1941	APRIL	200.	31.	3366.	88.68
1941	MAY	163.	0.	3332.	82.40
1941	JUNE	106.	79.	3187.	76.86
1941	JULY	0.	0.	3104.	80.08
1941	AUGUST	0.	0.	3069.	62.99
1941	SEPTEMBER	0.	0.	3012.	87.72
1941	OCTOBER	0.	24.	3082.	69.05
1941	NOVEMBER	40.	122.	3242.	77.34
1941	DECEMBER	200.	172.	3190.	87.28

		Z3	Z4	WATER LEVEL	PROBABILITY
1942	JANUARY	172.	61.	3133.	96.20
1942	FEBRUARY	24.	156.	3324.	70.79
1942	MARCH	200.	127.	3491. (+)	57.66
1942	APRIL	200.	154.	3598. (+)	12.74
1942	MAY	200.	67.	3576. (+)	9.21
1942	JUNE	200.	149.	3310.	40.00
1942	JULY	59.	0.	3194.	85.74
1942	AUGUST	0.	0.	3095.	81.53
1942	SEPTEMBER	0.	0.	3039.	81.06
1942	OCTOBER	0.	19.	3012.	76.90
1942	NOVEMBER	0.	34.	3035.	73.46
1942	DECEMBER	0.	47.	3096.	87.32
1943	JANUARY	29.	1.	3179.	96.20
1943	FEBRUARY	53.	144.	3305.	70.79
1943	MARCH	200.	106.	3093.	60.62
1943	APRIL	0.	0.	3088.	38.15
1943	MAY	0.	0.	3075.	34.71
1943	JUNE	0.	0.	3183.	26.88
1943	JULY	37.	0.	3178.	85.99
1943	AUGUST	20.	20.	3052.	81.95
1943	SEPTEMBER	0.	0.	3050.	44.39
1943	OCTOBER	0.	42.	3048.	81.86
1943	NOVEMBER	0.	85.	3196.	77.27
1943	DECEMBER	171.	120.	3153.	90.72
1944	JANUARY	122.	45.	3088.	96.20
1944	FEBRUARY	0.	111.	3201.	69.45
1944	MARCH	98.	71.	3344.	61.29
1944	APRIL	200.	60.	3155.	88.53
1944	MAY	0.	0.	3207.	73.24
1944	JUNE	0.	36.	3303.	76.49
1944	JULY	156.	0.	3115.	86.10
1944	AUGUST	0.	0.	3043.	77.91
1944	SEPTEMBER	0.	0.	2993.	60.29
1944	OCTOBER	0.	12.	3139.	59.20
1944	NOVEMBER	129.	151.	3235.	77.34
1944	DECEMBER	200.	188.	3289.	83.48
1945	JANUARY	200.	177.	3280.	91.76
1945	FEBRUARY	200.	191.	3389.	70.61
1945	MARCH	200.	150.	3404. (+)	35.50
1945	APRIL	200.	69.	3239.	82.96
1945	MAY	16.	1.	3187.	82.43
1945	JUNE	0.	0.	3108.	68.17
1945	JULY	0.	0.	3040.	31.79
1945	AUGUST	0.	0.	2941.	34.46
1945	SEPTEMBER	0.	0.	2913.	2.81
1945	OCTOBER	0.	8.	2932.	16.88
1945	NOVEMBER	0.	36.	3031.	56.57
1945	DECEMBER	0.	94.	3128.	90.56
1946	JANUARY	9.	8.	3187.	96.13
1946	FEBRUARY	8.	9.	3297.	51.19
1946	MARCH	9.	9.	3348.	54.57
1946	APRIL	9.	9.	3298.	21.09
1946	MAY	9.	6.	3214.	8.79
1946	JUNE	6.	0.	3157.	5.07
1946	JULY	0.	0.	3094.	11.64
1946	AUGUST	0.	0.	3001.	33.03
1946	SEPTEMBER	0.	0.	2949.	45.03
1946	OCTOBER	0.	0.	2956.	74.23
1946	NOVEMBER	0.	0.	3042.	59.54
1946	DECEMBER	0.	0.	3126.	75.57

		Z3	Z4	WATER LEVEL	PROBABILITY
1947	JANUARY	0.	0.	3213.	96.19
1947	FEBRUARY	0.	0.	3395.	53.64
1947	MARCH	0.	0.	3843. (+)	54.73
1947	APRIL	0.	0.	3960. (+)	79.00
1947	MAY	0.	0.	3943. (+)	32.30
1947	JUNE	0.	0.	3889. (+)	7.49
1947	JULY	0.	0.	3812. (+)	12.31
1947	AUGUST	0.	65.	3664. (+)	18.64
1947	SEPTEMBER	200.	120.	3379.	1.73
1947	OCTOBER	200.	106.	3174.	83.50
1947	NOVEMBER	96.	92.	3099.	77.34
1947	DECEMBER	12.	84.	3175.	90.72
1948	JANUARY	113.	5.	3172.	96.20
1948	FEBRUARY	50.	148.	3228.	70.79
1948	MARCH	116.	74.	3120.	61.29
1948	APRIL	0.	0.	3205.	58.97
1948	MAY	0.	0.	3186.	80.22
1948	JUNE	0.	0.	3203.	68.84
1948	JULY	0.	0.	3339.	85.07
1948	AUGUST	200.	5.	3098.	81.95
1948	SEPTEMBER	0.	0.	3068.	83.20
1948	OCTOBER	0.	63.	3114.	83.28
1948	NOVEMBER	68.	113.	3093.	77.34
1948	DECEMBER	27.	91.	3104.	90.72
1949	JANUARY	36.	3.	3153.	96.20
1949	FEBRUARY	12.	139.	3150.	70.79
1949	MARCH	0.	45.	3134.	61.24
1949	APRIL	0.	0.	3138.	60.86
1949	MAY	0.	0.	3179.	58.30
1949	JUNE	0.	12.	3106.	74.20
1949	JULY	0.	0.	3029.	36.44
1949	AUGUST	0.	0.	2954.	29.03
1949	SEPTEMBER	0.	0.	2870. (-)	8.73
1949	OCTOBER	0.	12.	2868. (-)	2.69
1949	NOVEMBER	0.	30.	3043.	18.30
1949	DECEMBER	32.	128.	3063.	90.72
1950	JANUARY	21.	39.	3157.	96.20
1950	FEBRUARY	30.	143.	3315.	70.79
1950	MARCH	200.	113.	3178.	59.82
1950	APRIL	0.	0.	3241.	87.05
1950	MAY	6.	0.	3202.	82.41
1950	JUNE	0.	0.	3047.	71.63
1950	JULY	0.	0.	2957.	3.19
1950	AUGUST	0.	0.	2899. (-)	4.42
1950	SEPTEMBER	0.	0.	2898. (-)	1.31
1950	OCTOBER	0.	21.	2965.	15.50
1950	NOVEMBER	0.	72.	3181.	73.61
1950	DECEMBER	200.	141.	3171.	90.72
1951	JANUARY	164.	73.	3146.	96.20
1951	FEBRUARY	53.	163.	3247.	70.79
1951	MARCH	156.	87.	3297.	61.29
1951	APRIL	153.	36.	3117.	88.72
1951	MAY	0.	0.	3229.	55.99
1951	JUNE	22.	77.	3375.	76.86
1951	JULY	200.	72.	3214.	79.08
1951	AUGUST	63.	31.	3116.	81.95
1951	SEPTEMBER	0.	0.	3134.	91.64
1951	OCTOBER	35.	114.	3062.	84.40
1951	NOVEMBER	0.	77.	3083.	77.31
1951	DECEMBER	0.	75.	3168.	90.65

		Z3	Z4	WATER LEVEL	PROBABILITY
1952	JANUARY	105.	5.	3177.	96.20
1952	FEBRUARY	56.	147.	3276.	70.79
1952	MARCH	188.	85.	3254.	61.29
1952	APRIL	91.	29.	3200.	88.72
1952	MAY	0.	0.	3151.	80.94
1952	JUNE	0.	0.	3113.	52.77
1952	JULY	0.	0.	2960.	45.98
1952	AUGUST	0.	0.	2832. (-)	3.99
1952	SEPTEMBER	0.	0.	2806. (-)	0.00
1952	OCTOBER	0.	0.	2939.	0.00
1952	NOVEMBER	0.	94.	3036.	72.80
1952	DECEMBER	9.	105.	3182.	90.72
1953	JANUARY	147.	36.	3145.	96.20
1953	FEBRUARY	35.	157.	3181.	70.79
1953	MARCH	51.	66.	3148.	61.29
1953	APRIL	0.	0.	3143.	74.44
1953	MAY	0.	0.	3183.	61.06
1953	JUNE	0.	15.	3205.	74.70
1953	JULY	2.	0.	3110.	85.70
1953	AUGUST	0.	0.	3059.	68.30
1953	SEPTEMBER	0.	0.	3000.	81.39
1953	OCTOBER	0.	16.	3008.	62.19
1953	NOVEMBER	0.	52.	2986.	74.67
1953	DECEMBER	0.	12.	3003.	65.02
1954	JANUARY	0.	0.	3076.	75.22
1954	FEBRUARY	0.	78.	3132.	66.33
1954	MARCH	0.	59.	3324.	61.29
1954	APRIL	170.	47.	3182.	88.72
1954	MAY	0.	0.	3357.	78.57
1954	JUNE	170.	91.	3202.	76.86
1954	JULY	10.	0.	3214.	86.11
1954	AUGUST	46.	2.	3114.	81.95
1954	SEPTEMBER	0.	0.	3083.	90.38
1954	OCTOBER	0.	72.	3104.	83.96
1954	NOVEMBER	48.	104.	3132.	77.34
1954	DECEMBER	76.	99.	3180.	90.72
1955	JANUARY	135.	25.	3166.	96.20
1955	FEBRUARY	55.	153.	3245.	70.79
1955	MARCH	146.	81.	3251.	61.29
1955	APRIL	81.	30.	3236.	88.72
1955	MAY	10.	0.	3236.	82.43
1955	JUNE	0.	30.	3190.	76.60
1955	JULY	0.	0.	3229.	75.02
1955	AUGUST	51.	0.	3259.	81.93
1955	SEPTEMBER	168.	66.	3117.	92.30
1955	OCTOBER	30.	142.	3227.	84.40
1955	NOVEMBER	200.	177.	3212.	76.88
1955	DECEMBER	200.	153.	3089.	89.97
1956	JANUARY	54.	46.	3123.	96.20
1956	FEBRUARY	0.	137.	3231.	70.75
1956	MARCH	122.	72.	3269.	61.29
1956	APRIL	102.	35.	3281.	88.72
1956	MAY	62.	0.	3282.	82.42
1956	JUNE	38.	64.	3258.	76.86
1956	JULY	52.	0.	3203.	85.75
1956	AUGUST	28.	2.	3100.	81.95
1956	SEPTEMBER	0.	0.	3015.	85.18
1956	OCTOBER	0.	8.	3037.	65.20
1956	NOVEMBER	0.	77.	3091.	76.88
1956	DECEMBER	24.	93.	3181.	90.72

		Z3	Z4	WATER LEVEL	PROBABILITY
1957	JANUARY	130.	18.	3097.	96.20
1957	FEBRUARY	0.	110.	3390.	69.59
1957	MARCH	200.	145.	3265.	33.55
1957	APRIL	60.	0.	3222.	88.70
1957	MAY	0.	0.	3234.	81.82
1957	JUNE	0.	34.	3178.	76.69
1957	JULY	0.	0.	3197.	69.06
1957	AUGUST	12.	0.	3117.	81.93
1957	SEPTEMBER	0.	0.	3131.	89.26
1957	OCTOBER	29.	105.	3076.	84.40
1957	NOVEMBER	10.	87.	3144.	77.34
1957	DECEMBER	81.	100.	3099.	90.72
1958	JANUARY	37.	11.	3162.	96.20
1958	FEBRUARY	26.	140.	3225.	70.79
1958	MARCH	105.	69.	3166.	61.29
1958	APRIL	0.	0.	3190.	84.50
1958	MAY	0.	0.	3130.	77.41
1958	JUNE	0.	0.	3219.	41.55
1958	JULY	54.	0.	3136.	85.69
1958	AUGUST	0.	0.	3053.	80.72
1958	SEPTEMBER	0.	0.	3048.	61.90
1958	OCTOBER	0.	47.	3029.	81.78
1958	NOVEMBER	0.	58.	3061.	76.47
1958	DECEMBER	0.	68.	3146.	90.26
1959	JANUARY	85.	7.	3159.	96.20
1959	FEBRUARY	32.	147.	3167.	70.79
1959	MARCH	23.	58.	3155.	61.29
1959	APRIL	0.	0.	3214.	76.17
1959	MAY	0.	0.	3223.	80.95
1959	JUNE	0.	19.	3305.	76.18
1959	JULY	142.	0.	3225.	85.94
1959	AUGUST	74.	15.	3096.	81.95
1959	SEPTEMBER	0.	0.	3027.	84.26
1959	OCTOBER	0.	21.	2987.	72.79
1959	NOVEMBER	0.	16.	3033.	65.57
1959	DECEMBER	0.	57.	3173.	88.28
1960	JANUARY	126.	20.	3174.	96.20
1960	FEBRUARY	68.	155.	3253.	70.79
1960	MARCH	160.	85.	3160.	61.29
1960	APRIL	0.	0.	3223.	84.94
1960	MAY	0.	0.	3256.	82.15
1960	JUNE	5.	57.	3171.	76.86
1960	JULY	0.	0.	3211.	61.20
1960	AUGUST	25.	0.	3129.	81.90
1960	SEPTEMBER	0.	0.	3122.	91.38
1960	OCTOBER	15.	104.	3248.	84.40
1960	NOVEMBER	200.	187.	3202.	76.09
1960	DECEMBER	200.	126.	3209.	90.70
1961	JANUARY	195.	59.	3135.	96.20
1961	FEBRUARY	39.	165.	3209.	70.79
1961	MARCH	100.	76.	3149.	61.29
1961	APRIL	0.	0.	3162.	78.30
1961	MAY	0.	0.	3212.	69.47
1961	JUNE	0.	38.	3220.	76.58
1961	JULY	8.	0.	3132.	85.64
1961	AUGUST	0.	0.	3022.	74.58
1961	SEPTEMBER	0.	0.	2952.	32.90
1961	OCTOBER	0.	0.	2941.	29.68
1961	NOVEMBER	0.	21.	3031.	51.68
1961	DECEMBER	0.	85.	3082.	90.27

		Z3	Z4	WATER LEVEL	PROBABILITY
1962	JANUARY	25.	17.	3173.	96.20
1962	FEBRUARY	47.	142.	3197.	70.79
1962	MARCH	66.	66.	3316.	61.29
1962	APRIL	159.	44.	3251.	88.72
1962	MAY	33.	0.	3199.	82.43
1962	JUNE	0.	0.	3149.	71.48
1962	JULY	0.	0.	3170.	58.56
1962	AUGUST	0.	0.	3055.	81.58
1962	SEPTEMBER	0.	0.	3013.	43.63
1962	OCTOBER	0.	9.	3012.	68.51
1962	NOVEMBER	0.	51.	3232.	75.10
1962	DECEMBER	200.	170.	3181.	88.90
1963	JANUARY	167.	67.	3205.	96.20
1963	FEBRUARY	125.	159.	3210.	70.79
1963	MARCH	102.	86.	3476. (+)	61.29
1963	APRIL	200.	163.	3450. (+)	18.92
1963	MAY	200.	51.	3259.	80.07
1963	JUNE	0.	38.	3242.	76.79
1963	JULY	8.	0.	3107.	85.13
1963	AUGUST	0.	0.	3114.	51.20
1963	SEPTEMBER	0.	73.	3210.	92.30
1963	OCTOBER	138.	137.	3147.	84.40
1963	NOVEMBER	137.	129.	3081.	77.34
1963	DECEMBER	31.	98.	3159.	90.72
1964	JANUARY	110.	21.	3091.	96.20
1964	FEBRUARY	0.	103.	3164.	68.93
1964	MARCH	42.	61.	3372.	61.29
1964	APRIL	200.	91.	3295.	86.60
1964	MAY	85.	0.	3286.	82.43
1964	JUNE	48.	67.	3237.	76.86
1964	JULY	20.	0.	3187.	85.71
1964	AUGUST	2.	0.	3154.	81.95
1964	SEPTEMBER	6.	45.	3146.	92.30
1964	OCTOBER	42.	111.	3274.	84.40
1964	NOVEMBER	200.	200.	3156.	70.11
1964	DECEMBER	127.	101.	3219.	90.72
1965	JANUARY	188.	38.	3195.	96.20
1965	FEBRUARY	112.	164.	3194.	70.79
1965	MARCH	79.	79.	3246.	61.29
1965	APRIL	66.	29.	3370.	88.72
1965	MAY	156.	0.	3407. (+)	82.38
1965	JUNE	200.	103.	3451. (+)	76.82
1965	JULY	200.	127.	3396.	18.60
1965	AUGUST	200.	102.	3327.	76.21
1965	SEPTEMBER	200.	110.	3199.	90.88
1965	OCTOBER	129.	161.	3085.	84.40
1965	NOVEMBER	46.	105.	3226.	77.34
1965	DECEMBER	200.	150.	3474. (+)	90.28
1966	JANUARY	200.	200.	3396.	0.00
1966	FEBRUARY	200.	200.	3398.	37.48
1966	MARCH	200.	133.	3349.	45.48
1966	APRIL	180.	20.	3314.	88.72
1966	MAY	98.	0.	3263.	82.41
1966	JUNE	8.	60.	3293.	76.86
1966	JULY	102.	0.	3280.	85.83
1966	AUGUST	132.	6.	3234.	81.95
1966	SEPTEMBER	143.	58.	3140.	92.30
1966	OCTOBER	59.	147.	3099.	84.40
1966	NOVEMBER	57.	106.	3282.	77.34
1966	DECEMBER	200.	200.	3285.	72.05

		Z3	Z4	WATER LEVEL	PROBABILITY
1967	JANUARY	200.	163.	3289.	94.38
1967	FEBRUARY	200.	188.	3273.	70.52
1967	MARCH	190.	99.	3225.	61.29
1967	APRIL	53.	21.	3314.	88.72
1967	MAY	96.	0.	3207.	82.41
1967	JUNE	0.	0.	3303.	72.86
1967	JULY	148.	0.	3100.	85.94
1967	AUGUST	0.	0.	2993.	72.95
1967	SEPTEMBER	0.	0.	3053.	15.55
1967	OCTOBER	0.	70.	3070.	83.66
1967	NOVEMBER	30.	106.	3073.	77.34
1967	DECEMBER	0.	83.	3113.	90.70
1968	JANUARY	43.	0.	3174.	96.20
1968	FEBRUARY	41.	141.	3218.	70.79
1968	MARCH	96.	68.	3172.	61.29
1968	APRIL	0.	0.	3182.	85.85
1968	MAY	0.	0.	3117.	75.28
1968	JUNE	0.	0.	3017.	34.75
1968	JULY	0.	0.	2877.(-)	3.64
1968	AUGUST	0.	0.	2947.	0.00
1968	SEPTEMBER	0.	6.	2968.	60.31
1968	OCTOBER	0.	59.	3008.	62.78
1968	NOVEMBER	0.	87.	3175.	76.82
1968	DECEMBER	167.	126.	3077.	90.72
1969	JANUARY	37.	40.	3179.	96.20
1969	FEBRUARY	64.	146.	3428.(+)	70.79
1969	MARCH	200.	152.	3462.(+)	23.48
1969	APRIL	200.	101.	3310.	56.84
1969	MAY	89.	2.	3234.	82.43
1969	JUNE	0.	33.	3318.	76.66
1969	JULY	158.	0.	3103.	85.96
1969	AUGUST	0.	0.	3073.	72.91
1969	SEPTEMBER	0.	0.	3084.	88.47
1969	OCTOBER	0.	98.	3104.	84.30
1969	NOVEMBER	55.	105.	3121.	77.34
1969	DECEMBER	62.	98.	3168.	90.72
1970	JANUARY	120.	22.	3160.	96.20
1970	FEBRUARY	42.	151.	3300.	70.79
1970	MARCH	200.	105.	3482.(+)	60.79
1970	APRIL	200.	158.	3505.(+)	13.90
1970	MAY	200.	76.	3340.	61.28
1970	JUNE	72.	57.	3250.	76.86
1970	JULY	18.	0.	3168.	85.32
1970	AUGUST	0.	0.	3217.	80.32
1970	SEPTEMBER	117.	77.	3095.	92.30
1970	OCTOBER	0.	125.	3107.	84.40
1970	NOVEMBER	52.	101.	3117.	77.34
1970	DECEMBER	52.	95.	3150.	90.72

FLOOD CONTROL RESERVOIR SYSTEM DESIGN USING STOCHASTIC PROGRAMMING

A. Prékopa
Technical University of Budapest
and
Hungarian Academy of Sciences

T. Szántai
Technical University of Budapest

Mathematically a natural river system is a rooted directed tree where the orientations of the edges coincide with the directions of the streamflows. Assume that in some of the river valleys it is possible to build reservoirs the purpose of which will be to retain the flood, once a year, say. The problem is to find optimal reservoir capacities by minimizing total building cost eventually plus a penalty, where a reliability type constraint, further lower and upper bounds for the capacities are prescribed. The solution of the obtained nonlinear programming problem is based on the supporting hyperplane method of Veinott combined with simulation of multivariate probability distributions. Numerical illustrations are given.

1. FORMULATION OF THE PROBLEM.

The problem we are dealing with has various applications. The most immediate among these is the application for flood control reservoir system design. Hence we shall use terms corresponding to this in the sequel.

In the theory of graphs the tree is usually defined as a connected undirected graph without circuit. When a special vertex has been designated and called the root of the tree then the tree is a rooted tree.

Taking into account the physical problem we are going to formulate it will be convenient for us to define rooted directed tree which arises from a rooted tree is such a way that a direction is assigned to every edge. The direction we assign are inductively given as follows: starting from the root, select the neighbouring vertices and assign to every edge thus obtained the direction showing in the direction of the root; then we start from the neighbouring vertices and do the same etc. In such a way we obtain a scheme as in *Figure 1*. This will be our mathematical model for a natural river system.

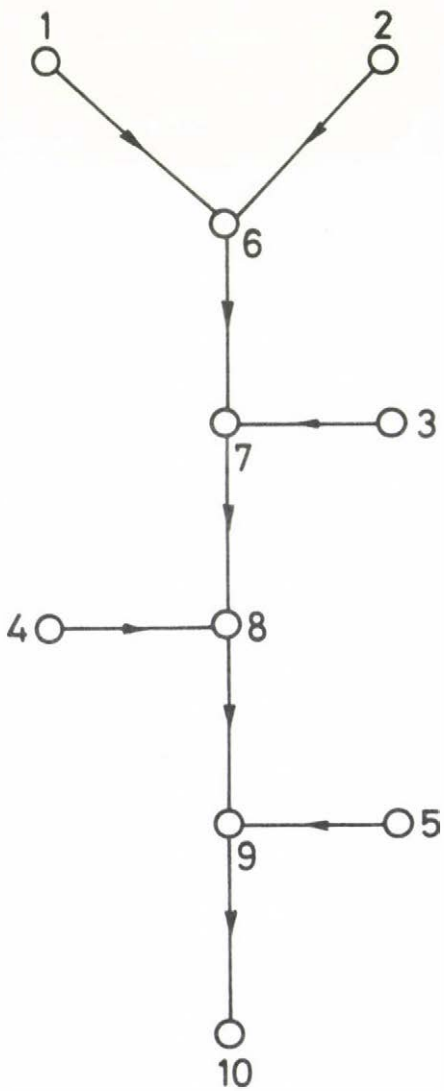


Fig. 1

Example for river system topology. Streamflow directions are indicated by arrows.

Assume that the number of vertices is $n+1$ and denote them by $a_1, \ldots, a_{n+1}$. The vertex a_i will be called an antecedent of the vertex a_j if a directed path leads from a_i to a_j. If a path of this kind exists then it is clearly unique. If this path consists of r edges, then we say that a_i is an antecedent of a_j of order r. Vertices without antecedents will be called terminal. Edges starting from terminal vertices will also be called terminal. Assume that the terminal vertices are $a_1, \ldots, a_m$ while a_{n+1} is the root of the tree. We assume further that there is only one edge going into a_{n+1} and this represents such a part of the river system where we can build reservoir. If there would be no possibility here for reservoir building, then we could split up the river system into subsystems and formulate at least two separate flood control problems. Let a_n denote that vertex with which the root is connected.

Assume now that we can build reservoirs in some parts of the river system, represented by some edges in our graph model, and the only purpose of the reservoirs will be retaining the flood. This we assume to exist periodically, once a year, say, and to be of a random character. The water comes from terminal points. We assume that the total water quantities can be separated in such a way that we can quantitatively give that amounts which can be lead between river banks and also that amounts which have to be retained by the reservoir system. Throughout the paper we are dealing with the water quantities to be retained only. When retaining the flood, we accept the following policy: first we start to fill up those upstream reservoirs which are located on terminal edges. Then, if the flood cannot be retained by these reservoirs, then start to fill up the reservoirs on the next edges by the overflown quantities and those input quantities which arrive on terminal edges without reservoir.

To every terminal vertex there corresponds a random input water quantity. These will be denoted by $x_1, \ldots, x_m$. The unknown reservoir capacities will be denoted by the symbols K_i. It is convenient to choose the subscript i in such a way that it coincide with the subscript of that vertex from which the edge, having the reservoir with capacity K_i, starts.

Let us define the quantities $x_{m+1}, \ldots, x_{n+1}$ recoursively as follows:

$$(1.1) \qquad x_j = \sum_{i \in A_j} [x_i - \min(a_i, K_i)] + \sum_{i \in B_j} x_i, \qquad j = m+1, \ldots, n+1$$

where $A_j (B_j)$ is the collection of those first order antecedents of the vertex a_j which are connected with a_j by an edge with (without) reservoir.

Now the reservoir system is capable to retain the flood if and only if the following condition holds:

$$(1.2) \qquad x_{n+1} = 0$$

or what is the same, the inequality holds:

$$(1.3) \qquad x_n \leq K_n.$$

Let us introduce the notation

$$(1.4) \qquad A = \bigcup_{j=m+1}^{n+1} A_j$$

For the design of the capacities K_i, $i \in A$ we formulate the following stochastic programming problem

$$\text{minimize} \quad \sum_{i \in A} [c_i(K_i) + E(\mu)]$$

subject to

$$P(x_n \leq K_n) \geq p, \tag{1.5}$$
$$0 \leq K_i \leq V_i, \quad i \in A$$

where p is a fixed probability, near unity in practice, the V_i, $i \in A$ are prescribed numerical upper bounds for the unknown capacities, E is the symbol of expectation and μ is a random penalty of the wrong deviation $x_n - K_n > 0$ provided it exists. Using lienar penalty, μ is defined in the following manner

$$\mu = \begin{cases} q(x_n - K_n) & \text{if} \quad x_n > K_n, \\ 0 & \text{otherwise}, \end{cases} \tag{1.6}$$

where q is a given nonnegative constant.

Variants of Problem (1.5) may also be of great practical interest. E.g. instead of retaining the flood by the total system of reservoirs we may prescribe the same for a number of subsystems and thus instead of the single inequality $x_n \leq K_n$ we may have a collection of such inequalities. Then either we prescribe a lower bound for the joint occurance of these inequalities or use separate probabilistic constraints.

The use of the joint probability distribution of the random variables $x_1, \ldots, x_m$ is fundamental in our reservoir system design methodology. Simple argument concerning simple system shows taht the reservoir capacities satisfying a prescribed reliability level strongly depend on the joint probabilistic behaviour of the input water quantities.

2. MATHEMATICAL PROPERTIES OF THE RESERVOIR SYSTEM DESIGN MODEL.

First we recall some fundamental mathematical concepts and theorems what we need in the sequel.

A nonnegative function f defined on R^k is said to be logarithmic concave (logconcave) if for every $x,y \in R^k$ and $0 < \lambda < 1$ we have the inequality

$$(2.1) \qquad f(\lambda x + (1-\lambda)y) \geq [f(x)]^{\lambda}[f(y)]^{1-\lambda}$$

If instead of (2.1) we have

$$(2.2) \qquad f(\lambda x + (1-\lambda)y) \geq \min[f(x),f(y)]$$

then f is said to be quasi-concave. In this latter case f need to be a nonnegative function. (2.1) obviously implies (2.2).

A probability measure defined on the measurable subsets of the space R^k is said to be logarithmic concave (logconcave) if for every pair A,B of convex subsets of R^k and every $0 < \lambda < 1$ we have the inequality

$$(2.3) \qquad P(\lambda A + (1-\lambda)B) \geq [P(A)]^{\lambda}[P(B)]^{1-\lambda}$$

If instead of (2.3) we have

$$(2.4) \qquad P(\lambda A + (1-\lambda)B) \geq \min[P(A),P(B)],$$

then P is called a quasi-concave probability measure.

Theorem 1. ([5],[7]).

If the probability measure P is absolutely continuous and is generated by a logconcave probability function, then P is a logconcave probability measure.

Theorem 2. ([2]).

If the probability measure P is absolutely continuous and is generated by the probability density f for which f is a convex function in the entire space R, then P is a quasi-concave probability measure.

Theorem 3. ([8]).

If a random vector y has logconcave (quasi-concave) probability distribution in R^q and $x = Ay + b$, where A is a constant $k \times q$ matrix and b is a constant k-component vector, then x has a logconcave (quasi-concave) probability distribution. (The proof is given for the logconcave distribution but the case of a quasi-concave probability distribution needs only trivial modification.)

Theorem 4.([6]).

Let $g_1(K,y)$,..., $g_r(K,y)$ be concave functions of all variables contained in the vectors K,y where $K \in R^k$ and $y \in R^q$. Assume that y is a random vector having logconcave (quasi-concave) probability distribution. Then

$$h(K) = P(g_i(K,y) \geq 0, \quad i = 1,\ldots,r) \tag{2.5}$$

is a logconcave (quasi-concave) function of the variable $K \in R^k$.

The most common example for logconcave probability distribution is the normal distribution. The nondegenerated normal distribution defined in R^k has the following density function

$$f(z) = \left[\frac{1}{(\det C)(2\pi)^n}\right]^{\frac{1}{2}} e^{-\frac{1}{2}(z-t)'C^{-1}(z-t)}, \quad z \in R^k \tag{2.6}$$

where C is a positive definite matrix, equal to the covariance matrix and t is a real constant vector equal to the expectation vector of the vector valued random variable, f is clearly a logconcave point function hence by *Theorem 1.* the corresponding probability distribution is logconcave.

A further example will be dealt with in the sequel is a special multivariate gamma distribution introduced in [10]. This is the probability distribution of the random vector x where

(2.7) $$x = Ay,$$

the random vector y has independent, standard gamma distributed components and A is a matrix of 0 and 1 entries. A continuous probability distirbution is called standard gamma distribution if it has the following type of probability density:

(2.8) $$\frac{1}{\Gamma(\nu)} z^{\nu-1} e^{-z} \quad \text{if} \quad z > 0$$

and zero if $z \leq 0$; ν is a positive constant. If $\nu \geq 1$ then this density function is logconcave. Assuming this property to hold for the components of y first we realize that y has a logconcave distribution (the independence of the components implies that the joint density is the product of the densities of the components) and by *Theorem 3* we derive that x has a logconcave distribution.

If some of the components of y have parameters smaller than 1, then the probability distribution of x may fail to be logconcave. However, the joint probability distribution function of the components of x is always a logconcave (point) function. We need somewhat more, therefore we prove

Theorem 5.

If the random vector y has independent, standard gamma distributed components and A_1 is a matrix with non-negative entries such that the product $A_1 y$ can be formed, then the probability distribution function of the random vectors $A_1 y$ i.e. the function

(2.9) $$P(A_1 y \leq z)$$

is logconcave in z in the entire space.

Proof. Let y_i be a component of y and denote by ν_i the parameter of its porbability distribution. If $\nu_i \geqq 1$ then we leave y_i in its original form. If on the other hand $\nu_i < 1$ then write

$$y_i = (y_i^{\nu_i})^{\frac{1}{\nu_i}} \tag{2.10}$$

and observe that $y_i^{\nu_i}$ has the following probability density function

$$\frac{d}{dz} P(y_i^{\nu_i} < z) = \frac{d}{dz} P(y_i < z^{\frac{1}{\nu_i}}) =$$

$$= \frac{d}{dz} \int_0^{z^{\frac{1}{\nu_i}}} \frac{1}{\Gamma(\nu_i)} t^{\nu_i - 1} e^{-t} dt = \frac{1}{\Gamma(\nu_i + 1)} e^{-z^{\frac{1}{\nu_i}}} \tag{2.11}$$

for $z > 0$ and zero for $z \leqq 0$. Since (2.11) is a logconcave probability density and by (2.10) the randon variable y_i is a convex function of the random variable $y_i^{\nu_i}$ having this logconcave density, *Theorem 4* implies that

$$P(A_1 y < z) = P(z - A_1 y \geqq 0)$$

is a logconcave function of the variable z. This completes the proof of *Theorem 5*.

Theorem 6.

Let y be a random vector having positive valued components and assume that the logarithms of the components have a joint normal distribution. Assume that A_1 is a matrix with nonnegative entries such that the product $A_1 y$ can be formed. Then the joint probability distribution of the components of $A_1 y$ is a logconcave (point) function in the entire space.

Proof. The proof is similar to the proof of *Theorem 5*. We only have to write

$$y_i = e^{\log y_i}$$

for every component y_i of y and repeat the argument applied in the previous proof.

The Gumbel distribution is a favoured distribution to describe the probabilistic behaviour of random extrema. As far as we know, however, no multivariate Gumbel distribution exists with correlated components. This limits the applicability of this distribution to our problem. The one-dimensional probability distribution function of a Gumbel distribution has the form

(2.12) $$e^{-\lambda e^{-\mu z}}, \quad -\infty < z < \infty,$$

where $\lambda > 0$, $\mu > 0$ are constants. Taking the derivative we obtain a logconcave (point) function hence this distribution is logconcave.

In our numerical examples the multivariate normal and the multivariate gamma distributions will be used.

Theorem 7.

Every x_j, $j \geq m+1$ is a convex function of those variables among $x_1, \ldots, x_m$ which belong to vertices antecedent relative to a_j and of all K_i variables which belong to edges connecting two such vertices at least one of which is antecedent relative to a_j. In particular, if x is the vector of components $x_1, \ldots, x_m$ and K is the vector of components K_i, $i \in A$, then

(2.13) $$g(K,x) = K_n - x_n$$

is a concave function of x and K.

Proof. First we remark that the following equality holds:

$$x_i - \min(x_i, K_i) = \max(0, x_i - K_i) \tag{2.14}$$

Substituting this into (1.1), a simple induction shows the validity of the theorem. We only have to refer to the fact that the maximum of two convex functions is again a convex function. This proves the theorem.

Theorem 8.

If $x_1, \ldots, x_m$ have a logconcave (quasi-concave) joint distribution, then the probability

$$h(K) = P(g(K,x) \geqq 0) \tag{2.15}$$

is a logconcave (quasi-concave) function of the variable K.

Proof. Theorem 8 is an immediate consequence of Theorem 4 and Theorem 7.

Theorem 9.

If x has a nondegenerated normal distribution then the function h(K) has continuous gradient at every K. If x has the above mentioned multivariate gamma distribution i.e. x=Ay where A and y satisfy the mentioned assumptions, then h(K) has continuous gradient except at most for the points where at least one component of K is zero.

Proof. The condition $x_n \leqq K_n$ can be expressed in terms of linear inequalities so that a number of partial sums of the random variables $x_1, \ldots, x_m$ are smaller then or equal to some partial sums of the variables K_i, $i \in A$. This system of inequalities arises in such a way that we substitute (2.14) into (1.1) and split up subsequently all inequalities of the type $\max(a,b) \leqq c$ into the inequalities $a \leqq c$, $b \leqq c$. The linear transformed of x thus obtained has a matrix what we

denote by B. This clearly has nonzero rows and every pair of its rows is a pair of linearly independent vectors.

Consider row first the case of the multivariate normal distribution. Since x has a nondegenerated normal distribution, it follows that any two components of Bx are linearly independent i.e. the absolute value of their correlation coefficient is smaller than 1. It is well-known that in this case the joint probability distribution function has continuous gradient in the entire space. Since the probability (*2.15*) can be expressed so that we put conveniently chosen partial sums of the components of K into the arguments of the joint distribution function of Bx, our assertion follows.

In case of the multivariate gamma distribution the argumentation is very similar. Since we cannot quarantee the differentiability of the joint distribution function of Bx only at such points where at least one of the components is zero, it follows that if all components of K are positive then the function (*2.15*) has continuous gradient at the point K. This proves the theorem.

Theorem 10.

The penalty function $E(\mu)$ in Problem (*1.15*) is convex no matter what kind of probability distribution x has. If x has a normal joint distribution, then $E(\mu)$ has continuous gradient at every K. If x has the multivariate gamma distribution described above then $E(\mu)$ has continuous gradient except at most for such K vectors where at least one component is zero.

Proof. The theorem can be proved easily on the basis of Theorem (*3.1*) in [9].

3. SOLUTION OF PROBLEM (1.5).

For the solution of Problem (1.5) we use the supporting hyperplane method of Veinott [11]. First we summarize this method then show how it applies to our problem. The following nonlinear programming problem will be considered

$$(3.1) \qquad \begin{aligned} &\text{minimize} \quad h_o(x) \\ &\text{subject to} \quad h_i(x) \geqq 0, \quad i = 1,\ldots,s. \end{aligned}$$

Assume that the following conditions hold:

Condition 1. There exists a bounded convex polyhedron K^1 such that

$$(3.2) \qquad \{x \mid h_i(x) \geqq 0, \quad i = 1,\ldots,s\} \subset K^1.$$

Condition 2. The functions $-h_o$, $h_1,\ldots,h_s$ are quasi-concave and have continuous gradient on K^1.

Condition 3. There exists a z^1 such that $h_i(z^1) > 0$, $i = 1,\ldots,s$.

The procedure consists of two phases.

Phase I. Find a vector z^1 satisfying Condition 3.

Phase II. We perform subsequent iterations where the r-th iteration consists of the following two steps.

Step 1. Solve the problem

$$(3.3) \qquad \begin{aligned} &\text{minimize} \quad h_o(x) \\ &\text{subject to} \quad x \in K^r \end{aligned}$$

where K^r is a bounded convex polyhedron. Let x^r be an optimal solution to this problem. If $h_i(x^r) \geqq 0$, $i = 1,\ldots,s$, then x^r is an optimal solution to Problem (3.1). Otherwise go to Step 2.

Step 2. Let λ^r be the largest $\lambda (0 \leq \lambda \leq 1)$ for which the following inequality holds

$$h_i(z^1 + \lambda(x^r - z_1)) \geq 0, \quad i = 1,\ldots,s.$$

This one-dimensional problem can be solved e.g. by Fibonacci search. Let

(3.4) $$y^r = z^1 + \lambda^r(x^r - z^1).$$

If $h_o(y^r) - h_o(x^r) \leq \varepsilon$, then y^r is an approximate solution of Problem (3.1). Otherwise select a subscript i_r for which the equality $h_i(y^r)=0$ holds. Let

(3.5) $$K^{r+1} = \{x | x \in K^r, \ \nabla h_{i_r}(y^r)(x-y^r) \geq 0 \}$$

and go to Step 1., using r+1 instead of r.

Figure 2. illustrates one iteration of the above procedure.

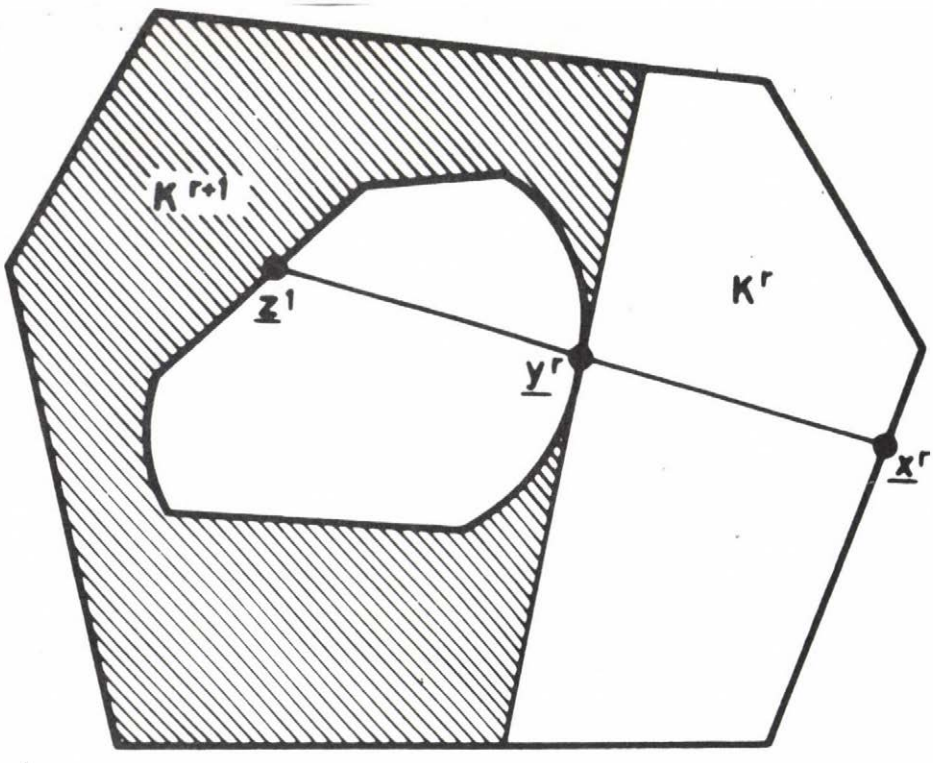


Fig. 2

An iterative step in the supporting hyperplane method of Veinott.

Veinott has shown that if instead of a fixed vector z^1 we use the following sequence

$$z^{r+1} = z^r + \beta(y^r - z^r) , \quad r = 1,2,... \tag{3.6}$$

in the subsequent iterations where $0 < \beta < 1$ is a fixed number, then his method reduces to the modified feasible direction method of Zoutendijk.

When solving Problem (1.5) we choose

$$K^1 = \{K \mid 0 \leq K_i \leq V_i , \quad i \in A\}.$$

We check whether $P(x_n < K_n) > p$ is satisfied in case of $K_i = V_i$, $i \in A$. This implies the fulfiment of Condition 3 and the vector of components $K_i = V_i$, $i \in A$ is selected to play the role of z^1. Thus only Phase II. had to be performed. The internal point was selected according to Relation (3.6) with $\beta = 0.5$. Unfortunately the sequence z^r, $r = 1,2,...$ converged very fast to a boundary point of the set determined by the probabilistic constraint and the procedure stopped before having obtained a good approxiamtion of the optimal solution. Tehrefore we used the modified relation

$$z^{r+1} = z^r + \frac{1}{r+1} (y^r - z^r) , \quad r = 1,2,... \tag{3.7}$$

which improved the method.

The values of the function $P(x_n \leq K_n)$ were obtained by simulation. Since simulation causes numerical inaccuracy, the following cautions were used.

a) In the course of the one-dimensional optimization we were seeking that λ for which the inequality

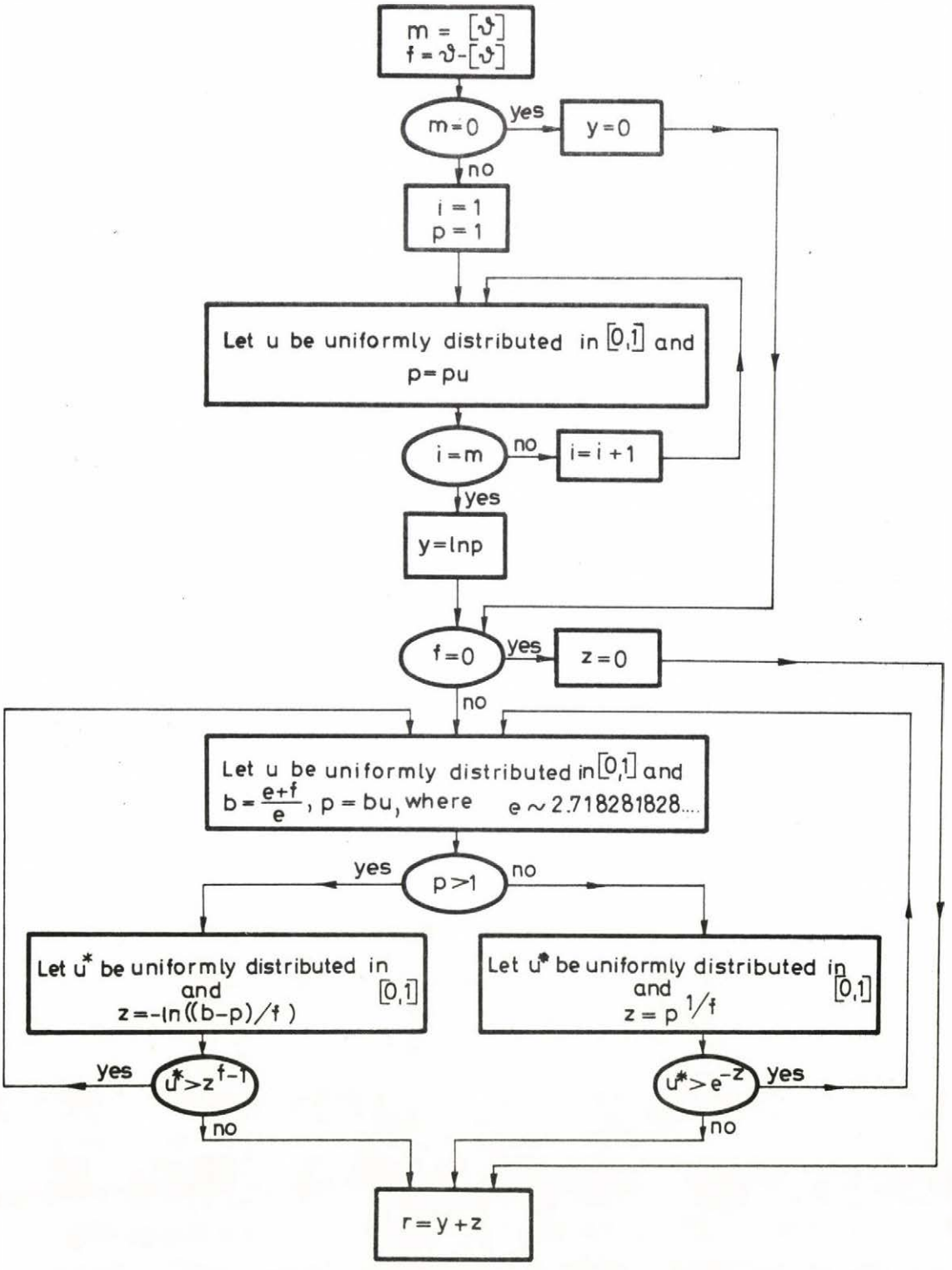
$$P(x_n \leq K_n) - p \geq -0.01$$

holds and the search procedure was stopped in case of an error smaller than 0.005. By this caution the procedure became somewhat slower but we prevented the new cut from cutting down feasible points from the constraining set of Problem (1.5). Primarily the inaccurate evaluation of the gradient causes such a danger. The one-dimensional optimization was a simple interval bisection.

b) Partial derivatives of the function $P(x_n \leq K_n)$ were computed numerically by the use of simulation corresponding to the values $K_i + 0,1$ and $K_i - 0,1$ for every $i \in A$ and the same random numbers were used to compute one gradient value.

The number of samples used for the simulation was 1000 when computing function values and 2000 when computing the derivatives of the probabilistic constraining function.

In case of the multivariate normal distribution the random numbers were generated by the use of a fast algorithm available on the UNIVAC 1108 computer of SCICON Ltd. on which the computations were performed. The method is due to Marsaglia and Bray [3]. The multigamma distributed random numbers were generated by the use of the method of Ahrens and Dieter [1]. The flow chart of this method is given by *Figure 3*. All programs were written in FORTRAN except for the program generating uniformly distributed random numbers and this latter was written in assambler.



Flow chart of the standard gamma distributed random number generation. Procedure of Ahrens and Dieter.

4. NUMERICAL EXAMPLES.

Consider the river system topology given by Figure 2, where the possible reservoir sites are also indicated. The variables defined by (*1.1*) are the following (x_1 , x_2 , x_3 , x_4 , x_5 are the terminal inputs):

$$
\begin{aligned}
x_6 &= x_1 - \min(x_1,K_1) + x_2 - \min(x_2,K_2),\\
x_7 &= x_3 - \min(x_3,K_3) + x_6,\\
x_8 &= x_4 + x_7,\\
x_9 &= x_8 - \min(x_8,K_8) + x_5 ,\\
x_{10} &= x_9 - \min(x_9,K_9).
\end{aligned}
\tag{4.1}
$$

Thus in terms of terminal variables and capacities Relation (*1.3*) is the following in this

$$
\begin{aligned}
x_9 = {} & x_1 - \min(x_1,K_1) + x_2 - \min(x_2,K_2) + x_3 - \min(x_3,K_3)\\
& + x_4 - \min[x_1-\min(x_1,K_1) + x_2 - \min(x_2,K_2) +\\
& + x_3 - \min(x_3,K_3) + x_4,K_8] + x_5 \leq K_9.
\end{aligned}
\tag{4.2}
$$

We shall not use penalty in our examples i.e. choose the penalty factor equal to zero. The problem is the following

$$
\begin{aligned}
&\text{minimize} \quad (0.4K_1 + 0.5K_2 + 0.6K_3 + 1.2K_8 + 1.8K_9)\\
&\text{subject to}\\
&\qquad P(x_9 \leq K_9) \geq p,\\
&\qquad 0 \leq K_1 \leq 1\\
&\qquad 0 \leq K_2 \leq 1\\
&\qquad 0 \leq K_3 \leq 1\\
&\qquad 0 \leq K_8 \leq 2\\
&\qquad 0 \leq K_9 \leq 3
\end{aligned}
\tag{4.3}
$$

where p will be chosen 0.8 and 0.9 in the following examples. Altogether 10 numerical examples are presented in Table 1. The following correlation matrices are used containing correlations of the random variables

$$R_1 = \begin{pmatrix} 1.0 & 0.0 & 0.6 & 0.4 & 0.0 \\ 0.0 & 1.0 & 0.5 & 0.3 & 0.3 \\ 0.6 & 0.5 & 1.0 & 0.7 & 0\ 6 \\ 0.4 & 0.3 & 0.7 & 1.0 & 0.4 \\ 0.0 & 0.3 & 0.6 & 0.4 & 1.0 \end{pmatrix},$$

$$R_2 = \begin{pmatrix} 1.0 & -0.5 & 0.0 & 0.3 & -0.5 \\ -0.5 & 1.0 & -0.8 & 0.0 & 0.2 \\ 0.0 & -0.8 & 1.0 & 0.0 & 0.3 \\ 0.3 & 0.0 & 0.0 & 1.0 & 0.0 \\ -0.5 & 0.2 & 0.3 & 0.0 & 1.0 \end{pmatrix},$$

$$R_3 = E,$$

where E is the 5×5 unit matrix. The expectations and standard deviations are the same in all examples. They are the following (expressed in a certain unit):

	Expectations	Standard deviations
x_1	0.8	0.2
x_2	1.5	0.3
x_3	1.2	0.6
x_4	0.5	0.4
x_5	0.7	0.3

For the case of the multigamma distribution we use the representation technique described in [10]. Of course only the case of R_1 has to be considered because R_2 contains negative entries too and the case of R_3 is trivial. For the case of R_1 we obtain the representation:

$$x_1 = \frac{1}{20}\ (y_1+y_2+y_3\qquad\qquad\qquad\qquad\qquad\qquad),$$

$$x_2 = \frac{3}{50}\ (\qquad y_4+y_5+y_6+y_7\qquad\qquad\qquad\qquad),$$

$$x_3 = \frac{3}{10}\ (y_1\qquad +y_4+y_5\qquad +y_8+y_9+y_{10}+y_{11}\qquad),$$

$$x_4 = \frac{8}{25}\ (\quad y_2\qquad\quad +y_6\quad +y_8+y_9\qquad +y_{12}\quad),$$

$$x_5 = \frac{9}{70}\ (\qquad y_4\qquad\quad +y_8\quad +y_{10}\qquad +y_{13}),$$

where $y_1,\ldots,y_{13}$ have standard gamma distributions with the following parameters:

$\nu_1 = 0.576$ $\qquad$ $\nu_8 = 0.14$

$\nu_2 = 0.16$ $\qquad$ $\nu_9 = 0.28$

$\nu_3 = 15.264$ $\qquad$ $\nu_{10} = 0.0498$

$\nu_4 = 0.315$ $\qquad$ $\nu_{11} = 2.055$

$\nu_5 = 0.585$ $\qquad$ $\nu_{12} = 0.7575$

$\nu_6 = 0.225$ $\qquad$ $\nu_{13} = 4.9404$

$\nu_7 = 23.875$

For the sake of completeness we give the system of inequalities containing partial sums of x_1 , x_2 , x_3 , x_4 , x_5 and K_1 , K_2 , K_3 , K_8 , K_9 which are together equivalent to the single inequality $x_9 \leq K_9$:

$$
\begin{array}{rcl}
x_5 & \leq & K_9, \\
x_4 & \leq & K_8+K_9, \\
x_1+x_4 & \leq & K_1+K_8+K_9, \\
x_2+x_4 & \leq & K_2+K_8+K_9, \\
x_3+x_4 & \leq & K_3+K_8+K_9, \\
x_1+x_2+x_4 & \leq & K_1+K_2+K_8+K_9, \\
x_1+x_3+x_4 & \leq & K_1+K_3+K_8+K_9, \\
x_2+x_3+x_4 & \leq & K_2+K_3+K_8+K_9, \\
x_1+x_2+x_3+x_4 & \leq & K_1+K_2+K_3+K_8+K_9.
\end{array}
$$

The optimal solutions, optimal objective function values, computing times and left hand side probability values at the end of the procedure are contained in Table 1.

Acknowledgement.

The authors express their thanks to Martin Beale who kindly offered the use of the UNIVAC 1108 computer of SCICON Ltd.

Table 1.

NUMERICAL RESULTS

Type of distribution	Correlation matrix	Probability level	K_1	K_2	K_3	K_8	K_9	Objective function	Computing time
MULTIVARIATE GAMMA	R_1	p=0.8	0.806735	1	1	1.355526	1.412243	5.591362	00:52:657
		p=0.9	0.751221	1	1	1.976085	1.398309	6.288746	00:35:688
	R_3	p=0.8	1	1	1	1.538713	1.103029	5.493909	00:16:785
		p=0.9	1	1	1	1.267790	1.848037	6.347815	00:11:343
MULTIVARIATE NORMAL	R_1	p=0.8	0.795523	1	1	1.590584	1.382698	5.815766	01:03:444
		p=0.9	0.997587	1	1	1.884778	1.524309	6.504525	00:25:126
	R_2	p=0.8	0.906312	1	1	1.350561	1.371008	5.551011	00:58:078
		p=0.9	0.833385	1	1	1.238889	1.830198	6.214377	00:51:426
	R_3	p=0.8	1	1	1	1.225805	1.430874	5.546541	00:43:461
		p=0.9	1	1	1	1.6499903	1.373814	5.952749	00:57:478

minutes
seconds
10^{-3} seconds

REFERENCES

[1] Ahrens, J.H. and Dieter,U., "Computer methods for sampling from gamma, beta, Poisson and binomial distributions", Computing 12 (1974) 223-246.

[2] Borell,C., "Convex set functions in d-space", Periodica Mathematica Hungarica 6(1975) 111-136.

[3] Marsaglia,G. and Bray,T.A., "A convenient method for generating normal variables", SIAM Rev. 6 (1964) 260-264.

[4] Ore,O., "Theory of graphs", American Mathematical Society Coll. Publ. 38(1962).

[5] Prékopa,A., "Logarithmic concave measures with application to stochastic programming", Acta Scientiarum Mathematicarum 32(1971) 301-316.

[6] Prékopa,A., "A class of stochastic programming decision problems", Mathematische Operationsforschung und Statistik 3(1972) 349-354.

[7] Prékopa,A., "On logarithmic concave measures and functions", Acta Scientiarum Mathematicarum 34(1973) 335-342.

[8] Prékopa,A., "Stochastic programming models for inventory control and reservoir system design", in: Inventory Control and Water Storage, Coll. Math. Soc. J. Bolyai 7, ed. A.Prékopa (North Holland Publ. Comp. Amsterdam-London, 1973) 229-245

[9] Prékopa,A., "Contributions to the theory of stochastic programming", Mathematical Programming 4(1973) 202-221.

[10] Prékopa,A., and Szántai,T., "A new multivariate gamma distribution and its fitting to empirical data", Water Resources Research to appear.

[11] Veinott,A.F., "The supporting hyperplane method for unimodal programming", Operations Research (1967) 147-152.

DYNAMIC TYPE STOCHASTIC PROGRAMMING MODELS

András Prékopa
Budapest

1. INTRODUCTION. Stochastic programming problems are formulated in such a way that first we start from a deterministic mathematical programming problem which will be called underlying deterministic problem, then observe that the model fails in the practice because some of the parameters in the problem are random variables, finally formulate another decision principle depending on the nature of the physical system and the randomness. The model which turns out is called stochastic programming model.

A stochastic programming model consists of one or more than one problems which are already deterministic, usually nonlinear. In contrast to this a problem, in which there are random parameters and it is to be solved (in principle) for every realization, will be called a random problem. Such problems will not be considered in this paper.

Mathematical programming problems will be subdivided here into two categories: to the first category belong those in which we only want to find a feasible solution whereas in problems belonging to the second category we have an objective function to be minimized or maximized.
In accordance with this, two kinds of uderlying problems will be considered. There is a further point, however, which has to be taken into account for practical reason. This will be clarified later on, now we only mention that this point doubles the number of underlying problems the number of which will therefore be altogether four. These are two couples. The first problem in the first couple is the following:

find $x_1, \ldots, x_{r+1}$

subject to the constraints

(1.1)

$$g_1(x_1) \geq 0,$$

$$g_2(x_1, x_2, \xi_1) \geq 0,$$

$$g_3(x_1, x_2, x_3, \xi_1, \xi_2) \geq 0,$$

$$\vdots$$

$$g_{r+1}(x_1, \ldots, x_{r+1} \xi, \ldots, \xi_r) \geq 0$$

$$h(x_1, \ldots, x_{r+1}) \geq 0.$$

In the second problem we have an objective function too.
The problem is the following:

maximize $f(x_1, \ldots, x_{r+1})$

subject to the constraints

$$g_1(x_1) \geq 0,$$

(1.2)

$$g_2(x_1, x_2, \xi_1) \geq 0,$$

$$g_3(x_1, x_2, x_3, \xi_1, \xi_2) \geq 0,$$

$$\vdots$$

$$g_{r+1}(x_1, \ldots, x_{r+1}, \xi_1, \ldots, \xi_r) \geq 0,$$

$$h(x_1, \ldots, x_{r+1}) \geq 0.$$

In the second couple of problems the number of decision variables is r.
The first problem is the following:

find $x_1, \ldots, x_r$

subject to the constraints

$$
\begin{aligned}
& g_1(x_1,\xi_1) \geqq 0, \\
& g_2(x_1,x_2,\xi_1,\xi_2) \geqq 0, \\
& \quad \vdots \\
& g_r(x_1, \ldots, x_r,\xi_1, \ldots,\xi_r) \geqq 0, \\
& h(x_1, \ldots, x_r) \geqq 0.
\end{aligned} \tag{1.3}
$$

In the second problem we again have objective function. The problem is the following:

maximize $f(x_1,\ldots, x_r)$

subject to the constraints

$$
\begin{aligned}
& g_1(x_1,\xi_1) \geqq 0, \\
& g_2(x_1,x_2,\xi_1,\xi_2) \geqq 0, \\
& \quad \vdots \\
& g_r(x_1,\ldots,x_r,\xi_1,\ldots,\xi_r) \geqq 0, \\
& h(x_1,\ldots,x_r) \geqq 0.
\end{aligned} \tag{1.4}
$$

In all problems the variables $x_1, \ldots, x_{r+1}$, the parameters $\xi_1,\ldots,\xi_r$ and the functions $g_1, \ldots, g_r$, h are supposed to be vector valued.

Our problems are supposed to be connected with a physical system which changes its states in time. Time is subdivided into periods the number of which is $r+1$ in case of the first couple and r in case of the second couple. $x_1, \ldots, x_r$ are decision variables, we decide on x_k just beginning of the k-th period. The random variables $\xi_1, \ldots, \xi_r$ are supposed to be observable in the subsequent periods. We observe ξ_k in the k-th period.

In case of Problems (1.1) and (1.2) the constraints have to be satisfied exactly if $\xi_1, \ldots, \xi_r$ are allowed to be random i.e. even in the stochastic programming problem whereas this will not be required in the stochastic programming problems constructed on the basis of Problems (1.3), (1.4).

The decision-observation scheme in case of Problems (*1.1*) and (*1.2*) is the following:

$$
\begin{array}{l}
\text{decision on } x_1, \\
\text{observation of } \xi_1 \\
\quad \vdots \\
\text{decision on } x_r, \\
\text{observation of } \xi_r, \\
\text{decision on } x_{r+1}.
\end{array}
\tag{1.5}
$$

First we choose x_1 so that $g_1(x_1) \geqq 0$, then observe ξ_1 and choose x_2 so that $g_2(x_1, x_2, \xi_1) \geqq 0$, etc.

The decision observation scheme in case of Problems (*1.3*) and (*1.4*) is the following:

$$
\begin{array}{l}
\text{decision on } x_1, \\
\text{observation of } \xi_1, \\
\quad \vdots \\
\text{decision on } x_r, \\
\text{observation of } \xi_r.
\end{array}
\tag{1.6}
$$

In this case the fulfilment of the constraints will not be required but random deviation will be penalized.

It should be mentioned that in practice frequently it is not possible to choose the decision vectors $x_1, \ldots, x_r$ so that we could exactly meet some equality constraints. If e.g. we are dealing with a water supply problem where large quantities are needed for irrigation, say, then in case of shortage it would be awkward to require a compensation because there may be no more water available. This explains that we in fact need both the above scheme couples.

In the stochastic programming model usually we have separate problem for the determination of each of the decision vectors. That problem the solution of which provides x_k will be called the k-th stage problem.

The purpose of this paper is to present a bunch of dynamic type or in other terms multi-stage stochastic programming models. Some of them are due to the author, others are from the literature. The main objective of the author is to give an insight into the mechanism of model construction and the structure of these models.
Primarily such models are included which are close to practical application. Technical details will be ommitted.

If the random variables $\xi_1, \ldots, \xi_r$ are independent, then the model formulation is relatively easy because the realized past history never influences the probability distribution of the random variables to be realized in the future. In practice, however, very frequently this is not the case. Moreover, time homogeneity would often be a strong restriction. Thus when formulating the stochastic programming models, it is advisable not to introduce such kinds of assumptions regarding the nature of randomness. This generality implies that the theoretically correct model formulations lead to very sophisticated problems the solution of which is very difficult if it is possible at all by the present technique.

The consequence of the above mentioned fact is that the model builder has to give up to a certain extent the statistical decision theoretic correctness in favour of the numerical solvability and the practical applicability. Though the practice requires an opportunistic behaviour in this respect, the author of this paper thinks that the principles listed below at least have to be taken into account when formulating dynamic type stochastic programming models.

A. Decide sequentially in time, after every observation of the random variables influencing the system.

B. At each decision use past history of the system i.e. base considerations on conditional distributions.

C. Take into account the probability distribution of the random variables which will be realized in the future.

D. Take into account the joint probability distribution of the random variables i.e. do not work only with the marginal distributions.

2. TWO-STAGE PROGRAMMING UNDER UNCERTAINTY.

This model was formulated by Dantzig ; and Madansky [6] and for a special case earlier by Dantzig [5] and Beale [1]. The underlying deterministic problem is of the type *(1.2)* where r=1. The decision variables x_1, x_2 will be denoted by x,y and in ξ_1 the subscript will be ommitted. The exact form of the underlying deterministic problem is the following:

$$
\begin{aligned}
& Ax \geqq b, \\
& Tx + My \geqq \xi, \\
& x \geqq 0, \quad y \geqq 0, \\
& \min(c'x + q'y),
\end{aligned}
\tag{2.1}
$$

where A, T, M are known matrices b, c, q are known vectors and ξ will be assumed to be random vector.

Accepting that ξ is random, we construct the stochastic programming model in the following way. First we fix that the set of feasible first stage decision vectors is the set of those x vectors for which there exists y such that the constraints of *(2.1)* are satisfied for every possible realization of the random vector ξ.

Assuming x,ξ to be fixed, we formulate the so called second stage problem as follows

$$
\begin{aligned}
& \text{minimize} \quad q'y \\
& \text{subject to the constraints} \\
& My \geqq \xi - Tx, \\
& y \geqq 0.
\end{aligned}
\tag{2.2}
$$

This provides an optimal y (we assume that it exists) depending on x and ξ and an optimal value $\mu = \mu(x,\xi)$.

The so called first stage problem fixes the value of the decision variable, x. This problem is the following:

$$
\begin{aligned}
&\text{minimize } [c'x + E(\mu(x,\xi))] \\
&\text{subject to the constraints} \\
&\qquad Ax \geqq b, \\
&\qquad x \geqq 0,
\end{aligned}
\tag{2.3}
$$

where E is the symbol of expectation.

If this model is applied in practice it is usually assumed that ξ is a discrete random vector. Denoting by $\xi^{(1)}, \ldots, \xi^{(N)}$ the possible values of ξ and by $p_1, \ldots, p_N$ the corresponding probabilities, Problem (2.3) can be rewritten in the following manner

$$
\begin{aligned}
&\text{minimize } (c'x + q'y^{(1)} + \ldots + q'y^{(N)}) \\
&\text{subject to the constraints} \\
&\begin{array}{lll}
Ax & & \geqq b, \\
Tx + My^{(1)} & & \geqq \xi^{(1)}, \\
Tx & + My^{(2)} & \geqq \xi^{(2)}, \\
\vdots & & \\
Tx & & + My^{(N)} \geqq \xi^{(N)}, \\
x \geqq 0, & y^{(1)} \geqq 0, & y^{(2)} \geqq 0, \ldots, y^{(N)} \geqq 0,
\end{array}
\end{aligned}
\tag{2.4}
$$

where $y^{(1)}, \ldots, y^{(N)}$ are new variables, the optimal $y^{(k)}$ is optimal solution of the second stage problem provided the realized value of ξ is $\xi^{(k)}$. Thus the optimal solution of Problem (2.4) provides at the same time the optimal solution of the first stage problem and the optimal solutions of the second stage problems for all realizations of ξ.

A considerable literature exists on this model. We refer to the book [8] in which some basic properties are discussed. An effective numerical solution technique is given in [21].

The random vector ξ may have a very wide set of possible values so that to require the solvability of the second stage problem for all realizations is frequently not possible. This is e.g. the case if we assume that ξ has a nondegenerated multivariate normal distribution. Then the second stage problem must have feasible solution for all right hand side vectors. This restricts very much the structure of the matrix M ([9]).

Now we present a variant of the two stage model: we require the solvability of the second stage problem only with a probability for which a lower bound is prescribed. To construct this problem, consider the following set

$$\{z \mid z \leqq My \ , \ y \geqq 0\}. \tag{2.5}$$

By the theorem of Weyl there exist vectors $d_1,\ldots, d_s$ such that z is an element of the set (2.5) if and only if it satisfies the inequalities

$$d_i' z \leqq 0, \quad i = 1 \ ,\ldots, s. \tag{2.6}$$

Thus x is a feasible solution of the first stage problem 2.3 if and only if

$$d_i'(\xi - Tx) \leqq 0, \quad i = 1,\ldots, s. \tag{2.7}$$

We shall prescribe a lower bound for the probability of the fulfiment of this inequality in the first stage problem.

The new second stage problem is defined in the following manner

$$\begin{aligned} &\text{minimize} && q'y + a'v \\ (2.8)' \qquad &\text{subject to the constraints} \\ & && My + v \geqq \xi - Tx. \\ & && y \geqq 0, \quad v \geqq 0, \end{aligned}$$

where ξ and x are fixed, a is some penalty vector having nonnegative components expressing a penalty for the unsolvability of Problem (2.2). Let $\mu(x)$ denote the random optimum value of Problem (2.8).

Then the new first stage problem can be formulated in the following manner:

$$\begin{aligned} &\text{minimize } [c'x + E(\mu(x))] \\ &\text{subject to the constraints} \\ (2.9) \qquad & P(d_i'x \leqq d_i' Tx, \quad i = 1, \ldots, s) \geqq p, \\ & Ax \geqq b, \\ & x \geqq 0. \end{aligned}$$

This variant of the two stage model was published in [12].

3. (r + 1)-STAGE PROGRAMMING UNDER UNCERTAINTY.

This is the generalization of the model described in the previous section. The model was published in [7]. The underlying deterministic problem is the following special case of Problem (1.2):

$$\begin{aligned} &\text{minimize } (q_1' x_1 + \ldots + q_{r+1}' x_{r+1}) \\ &\text{subject to the constraints} \\ & Ax_1 && \geqq b, \\ (3.1) \qquad & T_1 x_1 + T_2 x_2 && \geqq \xi_1, \\ & \vdots \\ & T_1 x_1 + \ldots + T_{r+1} x_{r+1} && \geqq \xi_1 + \ldots + \xi_r, \\ & x_1 \geqq 0, \ldots, x_{r+1} \geqq 0. \end{aligned}$$

The $(r + 1)$st stage problem is the following

$$\text{minimize } q'_{r+1} x_{r+1}$$

(3.2) subject to the constraints

$$T_1 x_1 + \ldots + T_{r+1} x_{r+1} \geq \xi_1 + \ldots + \xi_r,$$

$$x_{r+1} \geq 0,$$

where $\xi_1, \ldots, \xi_r$ are fixed at some realizations and $x_1, \ldots, x_r$ are considered parameters, also fixed for a moment. They are assumed to satisfy the first r constraints of Problem (3.2). The optimum value depends on $x_1, \ldots, x_r, \xi_1, \ldots, \xi_r$, let us denote it by

$$\mu_{r+1} = \mu_{r+1}(x_1, \ldots, x_r, \xi_1, \ldots, \xi_r).$$

The r-th stage problem is the following

$$\text{minimize } [q_r x_r + E_{\xi_1, \ldots, \xi_r}(\mu_{r+1})]$$

(3.3) subject to the constraints

$$T_1 x_1 + \ldots + T_r x_r \geq \xi_1 + \ldots + \xi_{r-1},$$

$$x_r \geq 0.$$

Here $\xi_1, \ldots, \xi_{r-1}$ are fixed at some realization and $x_1, \ldots, x_{r-1}$ are fixed parameters satisfying the first $r-1$ constraints of Problem (3.1).

Let $\mu_r = \mu_r(x_1, \ldots, x_{r-1}, \xi_1, \ldots \xi_{r-1})$ denote the optimum value of Problem (3.3). Proceeding on this way finally we formulate the first stage decision problem in the following manner

$$\text{minimize}[q'_1 x_1 + E_{\xi_1}(\mu_2) + \ldots + E_{\xi_1, \ldots, \xi_r}(\mu_{r+1})]$$

(3.4) subject to the constraints

$$Ax_1 \geq b,$$

$$x_1 \geq 0.$$

4. r - STAGE STOCHASTIC PROGRAMMING MODEL BASED ON THE UNDERLYING PROBLEM (*1.4*). This stochastic programming model [14] will be a sequence of problems, out of which exactly one has to be solved numerically to obtain the policy for one period.

Let g_{ij}, $j = 1,\ldots, m$ denote the j-th components of the vector valued function g_i and introduce the following notations

$$\mu_{ij} = \begin{cases} -g_{ij}(x_1,\ldots,x_i,\xi_1,\ldots,\xi_i) \text{ if } (g_{ij}(x_1,\ldots,x_i,\xi_1,\ldots\xi_i) \leq 0, \\ 0 \quad \text{otherwise,} \end{cases} \tag{4.1}$$

$$j = 1,\ldots,m_i \quad ; \quad i = 1,\ldots,r.$$

Assuming linear penalty function and denoting by the penalty for one unit of negative deviation of g_{ij} the expected value of the penalty for the r forthcoming period equals

$$\sum_{i=1}^{r} \sum_{j=1}^{m_i} E(q_{ij}\, \mu_{ij}). \tag{4.2}$$

The first problem among the mentioned problems is the following

$$\text{maximize } [f(x) - \sum_{i=1}^{r} \sum_{j=1}^{m_i} E(q_{ij}\mu_{ij})]$$

$$\text{subject to the constraints} \tag{4.3}$$

$$P(g_i(x_1,\ldots,x_i, \xi_1,\ldots,\xi_i) \geq 0, i = 1,\ldots,r) \geq p,$$

$$h_1(x) \geq 0,$$

where h_1 is the same function as h.

Out of the optimal solution of this problem we accept only x_1 as final value of the first decision variable. Keeping the notation x_1 for this fixed vector, we formulate the second problem in the following manner:

$$\text{maximize}\Big[f(x) - \sum_{i=2}^{r} \sum_{j=1}^{m_i} E(q_{ij}\mu_{ij})\Big]$$

(4.4) subject to the constraints

$$P(g_i(x_1,\ldots,x_i, \xi_1,\ldots, \xi_i) \geqq 0, \quad i = 2,\ldots,r) \geqq p,$$

$$h_2(x) \geqq 0.$$

We accept only x_2 as final decision vector etc. The last problem fixing x_r is the following

$$\text{maximize}\Big[f(x) - \sum_{j=1}^{m_r} E(q_{rj}\mu_{rj})\Big]$$

(4.5) subject to the constraints

$$P(g_r(x_1,\ldots,x_r, \xi_1,\ldots,\xi_r) \geqq 0) \geqq p,$$

$$h_r(x) \geqq 0.$$

An interesting simple special case of Problem (1.4) is the following production scheduling problem of one good:

$$\text{minimize } (c_1x_1 + \ldots + c_rx_r)$$

subject to the constraints

(4.6)
$$\begin{aligned} x_o + x_1 &\geqq \xi_1, \\ x_o + x_1 + x_2 &\geqq \xi_1 + \xi_2, \\ &\vdots \\ x_o + x_1 + \ldots + x_r &\geqq \xi_1 + \ldots + \xi_r, \end{aligned}$$

$$x_1 \geqq 0, \ldots, x_r \geqq 0,$$

where x_o denotes the initial stock, $x_1, \ldots, x_r$ the amounts to be produced in the subsequent time periods and $\xi_1, \ldots, \xi_r$ the demands finally $c_1, \ldots, c_r$ the production costs in the same periods. A more general scheme is the following:

$$\text{minimize} \quad (c_1' x_1 + \ldots + c_r' x_r)$$

subject to the constraints

$$\begin{aligned} Ax &\geq b \\ T_1 x_1 &\geq \xi_1, \\ T_1 x_1 + T_2 x_2 &\geq \xi_1 + \xi_2 \\ &\vdots \\ T_1 x_1 + \ldots + T_r x_r &\geq \xi_1 + \ldots + \xi_r, \end{aligned} \tag{4.7}$$

where x contains as components all components appearing in $x_1, \ldots, x_r$. This scheme differs from the underlying deterministic problem of the previous section only in that instead of equalities we use here inequalities.

A special case of the stochastic programming model system (4.3), (4.4), (4.5) was applied to obtain optimal operating policy for water reservoirs on the Tisza river in Hungary [14]. In this practical problem the functions $g_1, \ldots, g_r$ are nonlinear.

Nonlinear programming methods combined with simulation to obtain function and gradient values have been tested for this and various similar stochastic programming problems. Among these we mention the method of feasible directions [13] the SUMT method [19], the GRG method [11], the flexible tolerance method [14] and the cutting plane method [17].

Regarding the probability distributions the following two multivariate distributions were successfully applied: multivariate normal, multivariate gamma; for the special multivariate gamma distribution the reader is referred to [16].

5. SEQUENTIAL MAXIMIZATION OF PROBABILITY.

A general sequential probability maximization stochastic programming model system can be constructed on the basis of the underlying deterministic problem (1.3) similarly as we constructed the model system in the previous section on the basis of the underlying deterministic problem (1.4). A special case of this seems to be worth while to mention separately where we maximize the probability that a storage level remain between prescribed levels [15]. In the practical problem the water level of a lake is to be kept within prescribed limits where the water input is a discrete (in time) stochastic process and the water level can be controlled through a channel the capacity of which is given. Instead of water level we shall speak about water quantities in the sequel.

The underlying problem is the following:

$$\text{find } x_1, x_2, \ldots$$

$$\text{subject to the constraints}$$

$$(5.1) \qquad a_i \leq \xi_o + \xi_1 + \ldots + \xi_i - x_1 - \ldots - x_i \leq b_i,$$

$$0 \leq x_i \leq K, \qquad i = 1,2,\ldots,$$

where ξ_o is the initial water content of the lake, ξ_1, $\xi_2, \ldots$ are the monthly inflows and $x_1, x_2 \ldots$ are the monthly discharges through the channel the capacity of which is K.

Suppose now that $\xi_1, \xi_2, \ldots$ is a stochastic process and formulate the stochastic programming model.

In principle the water content regulation process is infinite. We formulate an infinitely long sequence of models; in every model we use conditional probability distribution where the whole past history of the system appears and we take into account the future N step ahead. Using months as periods, N = 12 seems to be a satisfactory long time during which stochastic dependence still exists. Let us introduce the notations

$$
\begin{aligned}
\xi_k &= \xi_o + \xi_1 + \ldots + \xi_k, \\
x_k &= \quad x_1 + \ldots + x_k,
\end{aligned}
\qquad k = 1,2,\ldots, \tag{5.2}
$$

where ξ_o is the initial water content of the lake. Let further $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ denote the prescribed lower, resp. upper bounds for the water quantities in the subsequent periods. The first stochastic programming problem is the following:

$$
\begin{aligned}
&\text{maximize} \quad P(a_k \leq \xi_k - X_k \leq b_k, \quad k = 1,\ldots,N) \\
&\text{subject to the constraints} \\
&0 \leq x_k \leq K, \qquad k = 1,\ldots,N.
\end{aligned}
\tag{5.3}
$$

We accept only x_1 as final value and formulate further problems in order to fix the other decision variables.

Before formulating the further problems we mention that there may be additional informations, beyond, the realized values of the process $\xi_1, \xi_2, \ldots$ which provide important informations for the solution of the practical problem. Such informations are e.g. the realized values of rainfall and meteorological data. It is customary to use all these informations for forecasting the input process and then the regulation problem turns out to be a deterministic problem if we accept the forecasted values as true values of the input process. Our procedure will be different. We use the available information as conditioning data and maximize conditional probability distribution subject to the constraints. In order to avoid sophisticated notations we assume that only one further stochastic process $\eta_1, \eta_2, \ldots$ provides information to our problem.

Assume now that we already solved n problems where the first one is Problem (5.3) and the random variables $\xi_1,\ldots,\xi_n, \eta_1,\ldots,\eta_n$ are realized.

Then the (n + 1) st problem is the following:

maximize $P(a_k \leq \xi_k - X_k \leq b_k, \quad k = n+1,\ldots, n+N \mid \xi_i, \eta_i, i=1,\ldots,n)$

subejct to the constraints

(5.4) $\qquad 0 \leq x_k \leq K, \quad k = n+1 ,\ldots, n+N.$

Out of the optimal $x_{n+1} ,\ldots, x_{n+N}$ we only accept x_{n+1} to be final decision value.

As regards the numerical solution of these problems we refer to the paper [15]. If we assume that the vector valued stochastic process (ξ_k, η_k) $k = 1,2,\ldots$ is Gaussian, then the conditional distribution needed in Problem (5.4) is also Gaussian hence the evaluation of the function values and the gradient values goes in the same way as those of absolute probabilities appearing in Problem (5.4).

The models of Section 4 and 5 are formulated theoretically in an incorrect way. They are, however, numerically evaluable hence practically applicable models.

Accepting the terms used in control theory we can say that the mentioned models are open loope models. A closed loop model for probability maximization will be described in the next section. There we use complete feedback but the problem is too complicated for numerical solution similarly as the problem of Section 3 and the multi-stage problem of Section 8.

6. CLOSED LOOP MODEL FOR TWO-STAGE MAXIMIZATION OF PROBABILITY

Out of the underlying Problem (4.1) we can formulate without difficulty theoretically correct models too in which feedback exist making the model - using control theoretic terms - closed loop model. In order to avoid complications first we restrict ourselves to the two stage case [14].

Let us start from the identity

$$P\begin{pmatrix} a_1 \leq \zeta_o + \xi_1 - x_1 & \leq b_1 \\ a_2 \leq \zeta_o + \xi_1 - \xi_2 + x_1 - x_2 & \leq b_2 \end{pmatrix} =$$

(6.1)

$$= \int_{a_1+x_1-\zeta_o}^{b_1+x_1-\zeta_o} P(a_2 \leq \zeta_o + \xi_1 + \xi_2 - x_1 - x_2 \leq b_2 \mid \xi_1) f_1(\xi_1) d\xi_1,$$

where f_1 is the probability density of the random variable ξ_1 (suppesed to be continuously distributed).

When deciding on x_2, the values of x_1 and ξ_1 are already fixed. Let us choose x_2 so that it maximizes the probability.

(6.2) $$P(a_2 \leq \zeta_o + \xi_1 + \xi_2 - x_1 - x_2 \leq b_2 \mid \xi_1)$$

for given x_1 and ξ_1. This is the second stage problem. Then optimal x_2 will be a function of two variables, $x_2 = x_2(x_1, \xi_1)$. Replacing this into the probability standing in the second row of (6.1), we get a function of the single variable x_1, and the maximization of this probability as a function of x_1, is the first stage problem. This can formally be written in the following way

$$(6.3)\quad \max_{0\leq x_1\leq K} \int_{a_1+x_1-\zeta_o}^{b_1+x_1-\zeta_o} \max_{0\leq x_2\leq K} P(a_2\leq \zeta_o+\xi_1+\xi_2-x_1-x_2\leq b_2 \mid \xi_1) f_1(\xi_1)\,d\xi_1 \,.$$

7. CLOSED LOOP MODEL FOR THE R-STAGE SEQUENTIAL MAXIMIZATION OF PROBABILITY. We now generalize the model of the previous section for the case of r stages. Assume that $\xi_1, \ldots, \xi_r$ have a continuous joint probability distribution and denote by

$$f_k(\xi_k \mid \xi_1, \ldots, \xi_{k-1}), \quad k = 2, \ldots, r$$

the conditional probability density of ξ_k given $\xi_1, \ldots, \xi_{k-1}$. Denote further by $f_1(\xi_1)$ the probability density of ξ_1. We hope that the reader will not be confused that the same notation is used for random variable as for argument of probability density function. It will turn out from the presentation that we want to simplyfy the description of the model. Let us define the following intervals

$$A_k(x_1, \ldots, x_k, \xi_1, \ldots, \xi_{k-1}) =$$

$$= \{\xi_k \mid a_k \leq \zeta_o+\xi_1+\ldots+\xi_{k-1}+\xi_k-x_1-\ldots-x_k\leq b_k, \quad k = 1, \ldots, r\}.$$

For simplicity we shall use the notation A_k too. By the theorem of total probability we have the identity

$$(7.1)\quad P\left(\begin{array}{lll} a_1 \leq & \zeta_o + \xi_1 - x_1 & \leq b_1 \\ a_2 \leq & \zeta_o + \xi_1 + \xi_2 - x_1 - x_2 & \leq b_2 \\ a_r \leq & \zeta_o + \xi_1 + \ldots + \xi_r - x_1 - \ldots - x_r & \leq b_r \end{array}\right) =$$

$$= \int_{A_1} \cdots \left(\int_{A_{r-1}} \left(\int_{A_r} f_r(\xi_r | \xi_1, \ldots, \xi_{r-1}) d\xi_r \right) f_{r-1}(\xi_{r-1} | \xi_1, \ldots, \xi_{r-2}) d\xi_{r-2} \right) \cdots f_1(\xi_1) d\xi_1$$

Now the rth stage problem is the following

$$\text{maximize} \int_{A_r} f_r(\xi_r \mid \xi_1, \ldots, \xi_{r-1}) d\xi_r$$

(7.2)

subject to the constraint

$$0 \leq x_r \leq K.$$

Here $x_1, \ldots, x_r, \xi_1, \ldots, \xi_{r-1}$ are fixed. The solution of this problem gives an

$$x_r = x_r(x_1, \ldots, x_{r-1}, \xi_1, \ldots, \xi_{r-1}),$$

where $x_1, \ldots, x_{r-1}$ are the previous policies and $\xi_1, \ldots, \xi_{r-1}$ the realized values of the same random variables. x_r is the rth stage optimal policy.

The (r-1)st stage problem is the following

$$\text{maximize} \int_{A_{r-1}} \left(\max_{0 \leq x_r \leq k} \int_{A_r} f_r(\xi_r | \xi_1, \ldots, \xi_{r-1}) d\xi_r \right) f_{r-1}$$

$$(\xi_{r-1} | \xi_1, \ldots, \xi_{r-2}) d\xi_{r-1}$$

(7.3) subject to the constraints

$$0 \leq x_{r-1} \leq K,$$

where the previously obtained optimal policy is inserted in the place of x_r producing there a dependence on x_{r-1}.

Proceeding on this way finally the first stage problem consists in the maximization with respect to the integral in the second row of (7.1) subject to the constraint $0 \leqq x_1 \leqq K$, where in the place of $x_2,\ldots,x_r$ the previously obtained optimal policies are inserted. This can formally be written so that before the integral sign where we integrate over A_k in (7.1) we put the word "max" and below the inequality $0 \leqq x_k \leqq K$.

8. THE E- AND THE P-MODELS.

In both cases the underlying deterministic problem is the following

$$
\begin{aligned}
&\text{minimize} \quad (c'x + q'y) \\
&\text{subject to the constraints} \\
&Ax \geqq \xi_1, \\
&Tx + My \geqq \xi_2.
\end{aligned}
\tag{8.1}
$$

Both models are two stage models, where x and y are the first, resp. second stage decision variables. Problem (8.1) is a special case of Problem (1.4) and the decision-observation scheme is given by (1.6) for the case $r = 2$.

Consider first the E-model [3]. In this case the second stage problem is the following

$$
\begin{aligned}
&\text{minimize} \quad q'y \\
&\text{subject to the constraints} \\
&P(Tx + My \geqq \xi_2 \mid \xi_1) \geqq p_2, \\
&y \geqq 0,
\end{aligned}
\tag{8.2}
$$

where p_2 is a prescribed probability $(0 < p_2 < 1)$. The optimal y depends on x and ξ_1. Let μ denote the random optimum value of Problem (8.2).

Now we formulate the first stage problem in the following manner:

$$\begin{aligned} &\text{minimize} \quad [c'x + E(\mu)] \\ &\text{subject to the constraints} \\ &\quad P(Ax \geqq \xi_1) \geqq p_1 , \\ &\quad x \geqq 0, \end{aligned} \tag{8.3}$$

where p_1 is a given probability $(0 < p_1 < 1)$.

The second stage problem of the P-model [4] is the same as that of the E-model. The first stage problem differs from Problem (8.3) in the formulation of the objective function.

The problem is the following:

$$\begin{aligned} &\text{minimize} \quad P(c'x + \mu \geqq K) \\ &\text{subject to the constraints} \\ &\quad P(Ax \geqq \xi_1) \geqq p_1, \\ &\quad x \geqq 0, \end{aligned} \tag{8.4}$$

where μ is again the random optimum value of the second stage problem and K is a given constant.

It is very difficult to give explicit expression for the expectation of μ as a function of x and ξ_1. It is good if we can avoid it and to solve the problem on some other way.

In an interesting special case, however, it was possible to give the required solutions for the r-stage problem [20]. The underlying deter-

ministic problem is Problem (*4.6*). Accepting that $\xi_1,\ldots,\xi_r$ are random variables and that the decision scheme is given by (*1.6*) formulate the stochastic programming model as follows. We write down the inequalities

$$
\begin{aligned}
&P(x_o + x_1 \geqq \xi_1) \geqq p_1 \\
&P(x_o + x_1 + x_2 \geqq \xi_1 + \xi_2 | \xi_1) \geqq p_2, \\
&\qquad\vdots \\
&P(x_o + x_1 + \ldots + x_r \; \xi_1 + \ldots + \xi_r | \xi_1,\ldots,\xi_{r-1}) \geqq p_r.
\end{aligned}
\tag{8.5}
$$

The rth stage problem will be:

$$
\begin{aligned}
&\text{minimize } c_r x_r \\
&\text{subject to the constraints} \\
&P(x_o + x_1 + \ldots + x_r \geqq \xi_1 + \ldots + \xi_1 / \xi_r \ldots, \xi_{r-1}) \geqq p_r \;,\; x_r \geqq 0,
\end{aligned}
\tag{8.6}
$$

where $x_1,\ldots,x_{r-1}$ are nonnegative parameters and $\xi_1,\ldots,\xi_{r-1}$ are fixed values. The (r-1) st stage problem is the following:

$$
\begin{aligned}
&\text{minimize } [c_{r-1} x_{r-1} + E\xi_{r-1}(c_r x_r)] \\
&\text{subject to the constraints} \\
&P(x_o + x_1 + \ldots + x_{r-1} \geqq \xi_1 + \ldots + \xi_{r-1} | \xi_1,\ldots,\xi_{r-2}) \geqq p_{r-1},
\end{aligned}
\tag{8.7}
$$

where $x_1,\ldots,x_{r-2}$ are nonnegative parameters, $\xi_1,\ldots,\xi_{r-2}$ are fixed values,

$$x_r = x_r(x_1,\ldots,x_{r-1},\xi_1,\ldots,\xi_{r-1})$$

is the optimal solution of Problem (*8.6*) given in parameteric form etc. The first stage problem is

$$
\begin{aligned}
&\text{minimize } [c_1 x_1 + E\xi_1(c_2 x_2) + \ldots + E\xi_1,\ldots,\xi_{r-1}(c_r x_r)] \\
&\text{subject to the constraints} \\
&P(x_o + x_1 \geqq \xi_1) \geqq p_1, \\
&\qquad x_1 \geqq 0.
\end{aligned}
\tag{8.8}
$$

The probabilistic inequalities in (8.5) can simply be converted ordinary linear inequalities. If the constants on the right hand side of the obtained inequalities form a nondecreasing sequence and $c_1 \leq \ldots \leq c_n$, then the optimal solutions of all problems can be given explicitely.

The formulation of the general E- and P-models are immediate on the basis of their description for the above special cases.

9. MULTI-PERIOD, MULTI-ITEM PRODUCTION PLANNING.

The model was published in [2]. Let r denote the number of periods and N the number of items. Introduce further the following notations:

a_{it}	amount of Product i sold in Period t,
p_{it}	amount of Product i produced in Period t,
s_{it}	amount of Product i in stock at, the end of Period t,
d_{it}	demand for product i in Period t.

The a_{it}, p_{it} are decision variables; the s_{it} are formally also decision variables which will be expressed recoursively; the d_{it} are constant in the underlying deterministic problem and are random variables in the stochastic programming problem.

To compose to objective function we need cost and price. Introduce the notations:

P_{it}	selling price of Product i in Period t,
C_{pit}	production cost of Product i in Period t,
C_{sit}	storage cost of Product i in Period t,
P	total productive capacity per time period.

The underlying deterministic problem is the following

$$\text{maximize} \quad \sum_{t=1}^{r} \sum_{i=1}^{N} (P_{it} a_{it} - C_{pit} p_{it} - C_{sit} s_{it})$$

subject to the constraints

$$
\begin{aligned}
& s_{i,t-1} + p_{it} - a_{it} = s_{it}, \\
& a_{it} \leq \min(s_{i,t-1} + p_{it},\ d_{it}), \\
& \sum_{i=1}^{N} p_{it} \leq P, \\
& p_{it} \geq 0, \quad a_{it} \geq 0, \\
& i = 1,\ldots,N; \quad t = 1,\ldots,r,
\end{aligned}
\tag{9.1}
$$

where s_{io} is the initial stock of Product i. This is a nonlinear programming problem, the nonlinearity is only in the second row in (9.1).

The second row in (9.1) can be rewritten in the following manner

$$
\begin{aligned}
a_{it} &\leq \min(s_{i,t-1} + p_{it},\ d_{it}) = \\
&= s_{i,t-1} + p_{it} + \min(0, d_{it} - s_{i,t-1} - p_{it}).
\end{aligned}
\tag{9.2}
$$

When formulating the stochastic programming problem, the random vectors $(d_{1t},\ldots,d_{Nt})$, $t = 1,\ldots,r$ are supposed to be independent. A special form of dependence is assumed between d_{it} and the sold amounts of the previous period. As a consequence of this, we do not have to work with conditional distributions.

The stochastic programming problem differs from Problem (9.1) only in the second row which reads now as

$$
\begin{aligned}
& a_{it} \leq s_{i,t-1} + p_{it} - E[\max(0, d_{it} - s_{i,t-1} - p_{it})] \\
& \qquad i = 1,\ldots,N; \quad t = 1,\ldots,r.
\end{aligned}
\tag{9.3}
$$

The expectation of right hand side is approximated by a linear expression, introducing new variables, where we assume that the d_{it} are normally distributed.

Using this model, we plan the a_{it} and p_{it} for a number of periods ahead without reconsideration from time to time the obtained results. Since the constraints hold only in expectation. Relation (9.3) can frequently be violated as the random d_{it} realize in time.

10. TWO SECTOR MULTI-STAGE ECONOMIC PLANNING.

The model was published in [22]. The underlying deterministic problem is the two sector economic model formulated in [10]. We shall use the following notations

I_t investment in Period t,

C_t consumption in Period t,

$Y_t = I_t + C_t$ national income in Period t,

$\lambda_t^{(I)} I_{t-1}$ proportion of the investment in Period t-1 devoted to new investment

$\lambda_t^{(C)} I_{t-1}$ proportion of the investment in Period t-1 devoted to consumption, where

$$\lambda_t^{(C)} + \lambda_t^{(I)} = 1,$$

$\beta_t^{(I)}$ rate of maximal investment growth (assumed to be linear)

$\beta_t^{(C)}$ of one unit of investment in Period t-1, rate of maximal consumption growth (assumed to be linear) of one unit of investment in Period t-1.

I upper bound for the sum of investments in all periods,

C lower bound for the consumption in every period.

The number of periods (years in practice) is r+1, their subscripts are $0,1,\ldots,r$. The I_t, C_t, $\lambda_t^{(I)}$, $\lambda_t^{(C)}$ are decision variables, the $\beta_t^{(I)}$, $\beta_t^{(C)}$, I, C are constants. Y_t is simply a notation. The deterministic problem is the following

11. GENERAL SCHEME FOR THE R-STAGE STOCHASTIC PROGRAMMING PROBLEM.

The scheme we are going to present below the most general model for the r-stage sequential decision problem.

Let $x_1,\dots,x_r$ be the decision variables as before and $\xi_1,\dots,\xi_r$ the random variables influencing the system in time. We do not require that these vectors have the same dimension, even the x vectors may have different dimensions among others and the same hodls for the ξ vectors. The decision-observation scheme is assumed to be of the type (*1.6*). Similar model can be constructed for the other case.

Let us introduce the following notations.

$P(k,\dots,r \mid \xi_1,\dots,\xi_{k-1})$ conditional joint probability distribution of $\xi_k,\dots,\xi_r$ given $\xi_1,\dots,\xi_{k-1}$, $k = 2,\dots,r$, joint probability

$P(1,\dots,r)$ distribution of the random variables $\xi_1,\dots,\xi_r$.

Now the kth stage decision vector x_k obviously has to be chosen so that it should be a function of the already realized values of $\xi_1,\dots,\xi_{k-1}$, the already fixed decision variables $x_1,\dots,x_{k-1}$ and the conditional distribution $P(k,..,r \mid \xi_1,\dots,\xi_{k-1})$, if $k \geq 2$. The first stage dicision vector is a function of the probability distribution $P(1,\dots,r)$. We assume further that $x_k \in D_k$ in the kth decision where $D_1,\dots,D_r$ are given, nonrandom sets. If $H_1,\dots,H_r$ denote these functions then we can write

$$x_1 = H_1(P(1,\dots,r)) \ , \quad X_1 \in D_1$$

$$x_k = H_k(P(k,\dots,r \mid \xi_1,\dots,\xi_{k-1})x_1,\dots,x_k,\xi_1,\dots,\xi_{k-1}) \tag{*11.1*}$$

$$x_k \in D_k \ , \ k = 2,\dots,r.$$

$$\text{maximize } Y_r = I_r + C_r$$

subject to the constraints

$$(10.1)\quad \begin{aligned} I_t - I_{t-1} &\leq \beta_t^{(I)} \lambda_t^{(I)} I_{t-1}, \quad t = 1,\ldots,r, \\ C_t - C_{t-1} &\leq \beta_t^{(C)} \lambda_t^{(C)} I_{t-1}, \quad t = 1,\ldots,r, \\ I_t &\geq 0, \quad C_t \geq C_o, \quad t = 0,1,\ldots,r \\ I_o + I_1 + \ldots + I_r &\leq I. \end{aligned}$$

In the original paper [10] equalities are instead of inequalities in the first two constraint groups.

Now we assume the $\beta_t^{(I)}$, $\beta_t^{(C)}$ to be random variables and construct the stochastic programming problem. The principle differs from those presented in the previous sections primarily in that by this problem we only want to determine the $\lambda_t^{(I)}$, $\lambda^{(C)}$.

These are to be determined so that the random optimum value of the random Program (*10.1*) reach a maximal expected value.

The random variables $\beta_t^{(I)}$, $\beta_t^{(C)}$, $t = 1,\ldots,r$ were supposed to be independent and having probability distribution of the type.

$$K(1 + \frac{x}{a})^p (1 - \frac{x}{b})^q$$

where K,a,b,p,q are constants, not necessarily the same for all random variables.

For the solution of the stochastic programming problem the method of simulation is proposed. Computations were performed in connection with a five year plan for India.

A generalization and a multisector version of the above model is described in [18].

The functions $H_1,\ldots,H_r$ are determined by a principle which consists in a problem of the following type.

$$\text{minimize } H(H_1,\ldots, H_r)$$

$$\text{subject to the constraints}$$

$$H_1 \in G_1 ,\ldots, H_r \in G_r$$

where H is a functional operating on $H_1,\ldots, H_r$ and $G_1,\ldots, G_r$ are sets to which $H_1,\ldots, H_r$ have to belong. They are partially given by the fact that $H_1,\ldots, H_r$ have to operate on the abovementioned probability distributions and given past history.

REFERENCES

[1] E.M.L. Beale, "On minimizing a convex function subject to linear inequalities", J. Roy. Stat. Soc. Ser.B., 17(1955) 173-184.

[2] E.M.L. Beale, J.J.H. Forrest and C.J. Taylor, "Multi-time period stchastic programming", in Proceedings of the International Conference on Stochastic Programming held in Oxford, England, 1974. Academic Press, to appear.

[3] A. Charnes, M.J.L. Kirby, "Optimal decision rules for the E-model of chance constrained programming". Cahiers du Centre d'Etudes de Recherche Operationelle 8(1966)5-44.

[4] A. Charnes, M.J.L. Kirby, "Some special P-models in chance constrained programming". Management Science 14 (1967) 183-195.

[5] G.B. Dantzig, "Linear programming under uncertainty" Management Science 1(1955) 197-206.

[6] G.B. Dantzig, A.Madansky, "On the solution of two stage linear programs under uncertainty" in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability (University of California Press, Berkeley, Calif. 1960) 165-176.

[7] G.B. Dantzig, "Linear programming and extensions", Princeton University Press, Princeton, N.J., 1963.

[8] P. Kall, "Stochastic Linear Programming", Econometrics and Operations Research (edited by W. Krelle and H.P. Künzi) 21, Springer Verlag, Berlin, New York, 1975.

[9] G. Kéri, "On the two stage programming under uncertainty" Studias Scientiarum Mathematicarum Hungarica, 5(1970) 37-40.

[10] P.C. Mahalanobis, "The approach of operational research to planning in India", Sankhya 16, (1955) 3-62.

[11] J. Mayer, "Computational experiences with the reduced gradient method", Colloquia Mathematica Societatis János Bolyai 12. Progress in Operations Research, Eger, Hungary 1974. (North Holland Publishing Company) 612-624.

[12] A. Prékopa, "Contributions to the theory of stochastic programming", Mathematical Programming 4(1973) 202-221.

[13] A. Prékopa, I. Deák, A. Ganczer and K. Patyi, "The STABIL stochastic programming model and its experimental application to the electrical energy sector of the Hungarian Economy" (in Hungarian)
Alkalmazott Matematikai Lapok 1/1975/ 3-22.
English translation to appear in Proceedings of the International Conference on Stochastic Programming held in Oxford (England), 1974. Academic Press.

[14] A. Prékopa, T. Szántai, "On multistage stochastic programming" Colloquia Mathematica Societatis János Bolyai 12. Progress in Operations Research, Eger (Hungary) 1974 (North Holland Publishing Company) pp. 733-755.

[15] A. Prékopa, "Optimal Control of a storage level using stochastic programming", Problems of Control and Information Theory 4(1975) 193-204.

[16] A. Prékopa, T. Szántai, "A new multivariate gamma distribution and its fitting to empirical data"
Alkalmazott Matematikai Lapok 1/1975/ to appear.
English translation to appear in Water Resources Research.

[17] A. Prékopa, T. Szántai, "Flood retaining models based on stochastic programming" , Water Resources Research, to appear.

[18] A. Prékopa, "Stochastic and dynamic planning models (in Hungarian), Institute of Economic Planning (Budapest) 1972.

[19] A. Prékopa, T. Rapcsák and I. Zsuffa, "A new method for serially linked reservoir system design"
Water Reseources Research, to appear.

[20] W.M. Raike, "Solutions to triangular conditional E-models arising from inventory problems", in Proceedings of the International Conference on Stochastic Programming held in Oxford (England), 1974. Academic Press. to appear.

[21] B. Strazicky, "On an algorithm for solution of the two stage stochastic programming problem", Operations Research Verfahren XIX, XX. Oberwolfach-Tagung über Operations Research 1973 (Verlag Anton Hain Meisenheim am Glan) 142-156.

[22] G. Tintner, N.S. Raghavan, "Stochastic linear programming applied to a dynamic planning model for India".